The universal measure of nonstochastic objects

Vladimir Vovk
Department of Computing, Security, and Mathematics
Royal Holloway, University of London

October 31, 2026

Abstract

The usual settings for discussing the notion of stochastic objects are very different from Kolmogorov's original structure-free setting, in which there is no predefined external structure on the objects (such as the lengths of binary strings). A natural modification of stochasticity, obtained by replacing finite sets as models with probability measures, is quasi-stochasticity. The structure-free setting is common in studies of quasi-stochasticity, and known results show that the universal machine generates non-quasi-stochastic objects with exponentially small probability. This paper shows that, on the other hand, the universal machine generates nonstochastic objects surprisingly often. It also systematically explores numerous connections between stochasticity and quasi-stochasticity in the structure-free setting. In particular, the definition of quasi-stochasticity can be modified in a natural way modelled on Kolmogorov's definition so that non-quasi-stochastic objects are produced as often as nonstochastic ones; the results for this modification imply all the results of this paper showing that nonstochastic objects are frequent.

Contents

1 Introduction 3

2 The universal measure of nonstochastic objects 5

3 Probability measures as models 7

4 A generalization of Theorem 1 10

5 Omnibus theorems 12

6 Conclusion 14

A Proofs 15

B Kolmogorov's questions 27

1 Introduction

The notion of stochastic objects was first introduced (in the currently recorded history) in Kolmogorov's talk at a seminar at Moscow State University on 26 November 1981. In the same talk he also posed the problem of estimating the abundance of stochastic and nonstochastic objects. (Alexander Shen's notes from that seminar have been published as [14, note 5.12].) Various solutions to this problem have been given by Shen in 1983 [15], Vladimir V'yugin in 1985 [20], and Nikolai Vereshchagin and Shen [18, Proposition 22]. See [16, Sect. 14.2] and [18, Sect. 2] for reviews.

A surprising feature of the existing literature on stochastic objects is that it always assumes some externally given nested structure on the objects whose stochasticity we are interested in.[1] In many cases the objects are the binary strings, with their lengths playing an important role. In other cases (e.g., in [15]), the objects are the nonnegative integers with the nested structure consisting of the sets $\{0,\dots,2^n-1\}$, $n=0,1,\dots$. Kolmogorov's original problem [14, note 5.12] is completely different: no externally given nested structure is used in measuring sets of nonstochastic objects. Kolmogorov's setting, discussed in Appendix B, still has an intrinsic nested structure (the sets of objects of complexity at most k), which the key results of this paper dispense with.

The standard problem of measuring the set of all (α,β)-nonstochastic objects (where α and β are parameters, to be defined in Sect. 2) has an important special case: do such objects even exist in a given level of the nested structure? (An example is the question of the existence of (α,β)-nonstochastic strings of length n.) Important results in this direction are due to Shen [15, Theorem 2] and Gács et al. [9, Theorem IV.2]. Such problems do not arise in the structure-free setting of this paper.

Different authors and publications use different ways of measuring sets of nonstochastic objects. Shen [15, Theorem 3] bounds the uniform probability measure of nonstochastic objects. Using the uniform probability measure for measuring the amount of nonstochastic objects appears somewhat unnatural, and [15, Theorem 4] covers non-uniform probability measures; Shen's Theorem 4, however, only gives a one-sided bound. V'yugin [20, Corollary and Theorem 3] bounds the universal semimeasure of nonstochastic objects (in our structure-free setting the universal semimeasure is a measure, but this is not necessarily true in structured settings). The existing result that is perhaps closest to the results of this paper about stochasticity is Theorem 251 in [16, Sect. 14.2]; this result uses the same version of the universal semimeasure as this paper (different from the one used by V'yugin). In the last two of these results, those in [20, 16], and in a plethora of related results in the literature, the bounds on the universal semimeasure of the (α,β)-nonstochastic objects are governed by the term $2^{-\alpha}$. Kolmogorov's setting described in Appendix B is very different, but we still have a similar prominent role for $2^{-\alpha}$. The nature of the first key result of this paper, Theorem 1, is even more different, and $2^{-\alpha}$ disappears.

One interpretation of the existing results on the universal semimeasure of nonstochastic objects is that "the probability of obtaining a non-stochastic object in a random process is negligible" [18, Sect. 4.6]. According to this paper's results, whether this is true in the structure-free setting depends on the details of the definition of stochasticity (and happens to be true for Vereshchagin and Shen's official definition [18, Definition 1]). If the definition uses finite sets as models, the situation in the structure-free setting is the opposite, and nonstochastic objects are generated extremely often by the universal semimeasure (Theorem 1 below). But if the definition uses probability measures as models, their counterparts become rare under the standard definitions (Theorem 7 below). In the terminology of [15, 9] and this paper, nonstochastic objects are frequent while non-quasi-stochastic

[1] In this paper I will follow Shen [15] and Gács et al. [9] and make a terminological distinction between stochastic objects, which are defined using finite sets as models, and quasi-stochastic objects, which are defined using probability measures as models.

objects are rare.

The notions of stochasticity and quasi-stochasticity are regarded as close in the standard structured settings (see, e.g., [15, Theorem 1] and [18, Proposition 2]). Therefore, nowadays the same term, "stochasticity", is often used for both stochasticity and quasi-stochasticity. In the structure-free setting a chasm opens up between stochasticity and quasi-stochasticity.

The first key result of this paper (Theorem 1) is stated in Sect. 2; it evaluates the likelihood of the universal machine producing nonstochastic objects. The result is very simple and implies that nonstochastic objects are produced very often. Another result in that section, Theorem 6, generalizes Theorem 1 and shows that the conditional universal measure of (α, β)-nonstochastic objects among objects of high complexity is bounded away from zero; see (8).

In the case of quasi-stochastic objects some interesting structure-free results exist in the literature, and they are in stark contrast with the results of this paper for stochastic objects. The most explicit ones originate from Epstein and Levin's work [6, Proposition 1]; for proofs, see [10, Proposition 1] and [5, Lemma 10]. (And the idea behind those results may go back as far as 1970 [3, the end of Sect. 1].) The report by Bauwens [2], posted slightly before Epstein and Levin's report [6], is somewhat less explicit. Results of this kind, demonstrating the paucity of non-quasi-stochastic objects, will be discussed in Sect. 3. I learned of the structure-free results by Epstein, Levin, and Bauwens, and of their relevance to quasi-stochasticity, with the help of AI tools.

The reason for the contrast between the abundance of nonstochastic objects and the paucity of non-quasi-stochastic objects is that the complexity of stochastic objects is simple (i.e., of low complexity), while this is not true for quasi-stochastic objects. The simplicity of complexity of stochastic objects was noted by Gács et al. [9, Lemmas III.8 and III.10], and in this paper it is replaced by the closeness of the prefix and plain complexities for stochastic objects (and further generalized in Sect. 4). Our second key result, Theorem 7 in Sect. 3, is a version of Epstein and Levin's results that formalizes the paucity of non-quasi-stochastic objects in a setting close to Kolmogorov's.

Stronger versions of Theorem 1 and complementary results are established in Sect. 4. Section 5 gives a more systematic picture covering two complexities (prefix and plain) and four randomness deficiencies (based on prefix and plain complexities and on universal e-values and p-values, to be defined there). I will state two omnibus results, one covering quasi-stochasticity (Theorem 13) and the other covering stochasticity (Theorem 14). These omnibus theorems contain the two key results mentioned earlier, the first one in a stronger form.

All nontrivial proofs for the statements in the main part of the paper are collected in Appendix A. Most of the proofs in this paper have been drafted using AI tools. At the beginning of Appendix A I will also briefly discuss the main techniques used in this paper.

Kolmogorov's original questions asked in 1981 concern the number of nonstochastic objects of complexity not exceeding a given level k. His specific questions are answered in Appendix B.

Let $\mathbb{N} := \{1, 2, \dots\}$ be the natural numbers without 0, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ be the natural numbers including 0, $[m] := \{1, \dots, m\}$ for each $m \in \mathbb{N}$, log be binary logarithm, C be plain (Kolmogorov) complexity [16, Sect. 1.1], K be prefix complexity [16, Sect. 4.3], and $\mathbf{m}$ be a fixed universal semimeasure on $\mathbb{N}$ [16, Sect. 4.2]. For a subset $A \subseteq \mathbb{N}$, we set $\mathbf{m}(A) := \sum_{a \in A} \mathbf{m}(a)$ (therefore, $\mathbf{m}(\{a\}) = \mathbf{m}(a)$ for all $a \in \mathbb{N}$). This makes $\mathbf{m}$ a measure, which is why we refer to it, above and below, as the *universal measure*. The words "increasing" and "decreasing" are used in the wide sense allowing intervals of constancy. A "binary string" means an element of $\{0,1\}^*$. For $x \in \{0,1\}^*$, $|x|$ is the length of x; $\Box$ is the empty string, $|\Box| = 0$.

In this paper we take $\mathcal{X}$ to be our basic space of finite objects whose (α, β)-stochasticity we are interested in. In all formal statements $\mathcal{X}$ can be replaced by $\mathbb{N}$ (but in informal discussion using $\mathbb{N}$ in place of $\mathcal{X}$ can cause confusion, since natural numbers play many other roles in this paper). The

basic space $\mathcal{X}$ can be any set in computable bijection with $\mathbb{N}$, such as the binary strings $\{0,1\}^*$. We will often refer to elements of the basic space $\mathcal{X}$ as objects and will usually denote them by ω. The complexities $C(\omega)$ and $K(\omega)$ and the universal measure $\mathbf{m}(A)$ for $A \subseteq \mathcal{X}$ are defined, naturally, as the corresponding functions (C, K, or $\mathbf{m}$) applied to the images of ω or A under the computable bijection. When we talk about a *shortest* plain or self-delimiting description of an object ω, we mean descriptions in the sense of the optimal decoders that determine C or K, respectively; therefore, shortest descriptions of ω have length $C(\omega)$ for plain descriptions and $K(\omega)$ for self-delimiting ones. When convenient, we identify $\mathcal{X}$ with $\mathbb{N}$ via the computable bijection, which gives a meaning to expressions such as $[m] \subseteq \mathcal{X}$ and $\max \Omega$ for finite $\Omega \subseteq \mathcal{X}$; in the appendices, we always do so. For concreteness, the reader can interpret $\mathcal{X}$ as $\mathbb{N}$ or $\{0,1\}^*$ in the main part of the paper.

For a brief summary of important results about finitary randomness, including connections with universal p-values and e-values, see [19, Appendix 1]; that summary only covers randomness w.r.t. uniform measures on finite sets (and so is relevant to stochasticity but not quasi-stochasticity).

2 The universal measure of nonstochastic objects

We say that an object $\omega \in \mathcal{X}$ is (α, β)-*stochastic*, where α is a nonnegative integer and $\beta \in [0, \infty]$, if there exists a finite set $\Omega \subseteq \mathcal{X}$ for which

- $\omega \in \Omega$;
- $K(\Omega) \le \alpha$;
- $K(\omega \mid \Omega) \ge \log |\Omega| - \beta$.

Otherwise ω is (α, β)-*nonstochastic*. In our notation we follow Kolmogorov (see Appendix B).

Let $\mathcal{N}_{\alpha,\beta} \subseteq \mathcal{X}$ be the set of all (α, β)-nonstochastic objects. We are interested in the asymptotics of $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ as $\alpha, \beta \to \infty$. Let

$$\tau(t) := \mathbf{m}(\{n \in \mathbb{N} \mid n > t\})$$

be the tail of $\mathbf{m}$; we allow $t \in [0, \infty)$ but are usually interested in $t \in \mathbb{N}_0$. Set $\tau(\infty) := 0$.

Theorem 1. *There exists $c > 1$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \le 2^{2^\alpha}$,*

$$\tau(\alpha)/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}) \le c\tau(\alpha). \tag{1}$$

The condition $\beta \le 2^{2^\alpha}$ in the statement of the theorem can be replaced by $\beta \le g(\alpha)$ for a fixed computable function g ("fixed" in the sense of c depending on g). Moreover, this condition is only needed for the lower bound. An even stronger version of Theorem 1 is

$$\tau(\max(\alpha, \beta))/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}) \le c\tau(\alpha); \tag{2}$$

cf. Lemma 3 below. In Sect. A.2 we will prove Theorem 1 in the form (2). The proof will show that the theorem continues to hold with $\mathcal{N}_{\alpha,\beta}$ replaced by its modification $\mathcal{N}^C_{\alpha,\beta}$ that has C in place of both occurrences of K in the definition of $\mathcal{N}_{\alpha,\beta}$.

The theorem says that $\mathbf{m}(\mathcal{N}_{\alpha,\beta}) = \Theta(\tau(\alpha))$ under mild conditions on β. Informally, nonstochastic objects are generated with probability that decays extremely slowly. The function $n \mapsto \tau(n)$ monotonically decreases to 0 but does so more slowly than any computable function monotonically decreasing to 0, such as $1/\log\log n$ (and we can take any number of logarithms). Let us state this well-known observation formally.

Lemma 2. *Let $f : \mathbb{N} \to (0, \infty)$ be a computable decreasing function such that $f(n) \to 0$ as $n \to \infty$. Then $\tau(n)/f(n) \to \infty$ as $n \to \infty$.*

Proof. Redefine $f := \min(f, 1)$. Set $h(t) := \sqrt{f(t)}$ for $t \in \mathbb{N}$ and $h(0) := 1$. Since f is decreasing and $f(t) \to 0$, $P(n) := h(n-1) - h(n)$, $n \in \mathbb{N}$, is a computable probability measure on $\mathbb{N}$ with tails $P(\{n \mid n > t\}) = h(t)$ for all $t \geq 0$. The universality of $\mathbf{m}$ gives $\mathbf{m} \geq cP$ for some $c > 0$, and so, as $t \to \infty$,

$$\frac{\tau(t)}{f(t)} \geq \frac{ch(t)}{f(t)} = \frac{c}{\sqrt{f(t)}} \to \infty. \qquad \square$$

Let us discuss in greater detail what appears to be the main difference between Kolmogorov's approach to stochasticity and the standard one relying on a given structure on the objects. Kolmogorov's definitions and questions, described in detail in Appendix B, are structure-free in that they only involve the complexities of the objects (the complexities can be, e.g., plain or prefix, conditional or unconditional) and other objects derived from the given objects (such as finite sets composed of the given objects). Complexities can be replaced by randomness deficiencies (which do not have to be defined via complexities); the point is that structure-free statements should involve only functions that are invariant, to within an additive constant, w.r.t. computable permutations of the objects. An example of a non-invariant function is provided by the length of a binary string. In this paper we will be interested only in structure-free statements. Theorem 1 is an example, and we will see numerous further ones later.

Now let us briefly discuss basic properties of τ and relations between τ and other well-known functions. The following lemma shows how insensitive τ is to the value of its argument, and it also shows that (2) is indeed a stronger version of Theorem 1.

Lemma 3. *Let $g : \mathbb{N}_0 \to \mathbb{N}_0$ be a computable increasing function. There is $c_g > 0$ such that $\tau(g(t)) \geq c_g \tau(t)$ for all $t \in \mathbb{N}_0$.*

Set

$$\sigma(t) := \mathbf{m}(\{n \mid K(n) > t\});$$

we will need this function for the interpretation of Theorem 6 later in this section. The following relation is well known, but I have not seen it stated explicitly in print.

Lemma 4. *For some constant $c > 1$,*

$$\tau/c \leq \sigma \leq c\tau. \tag{3}$$

Another interesting connection is with Chaitin's [4, Definition 5.3] version

$$\mathrm{BB}(s) := \max\{n \in \mathbb{N} \mid K(n) \leq s\} \tag{4}$$

(with $\max \emptyset := 0$) of the busy beaver function (see also [16, Sect. 1.2.2]). Let us set

$$\mu(t) := \min\{K(n) \mid n > t\}, \tag{5}$$

so that μ is, essentially, the inverse to the busy beaver function. Andreev [1, Theorem 1] proves that

$$\mu - O(\log \mu) \leq -\log \tau \leq \mu + O(1). \tag{6}$$

(Andreev's statement uses inverse functions to μ and $-\log \tau$, and in [1, Theorem 2] he shows the optimality of (6), crediting Gács [7] with part of his Theorem 2.)

Lemmas 3 and 4 will be proved in Sect. A.1.

The condition on β in Theorem 1 is weak, but it cannot be dropped. The following theorem describes the dependence of $-\log \mathbf{m}(\mathcal{N}_{\alpha,\beta})$ on β completely, albeit only with logarithmic precision in α.

Theorem 5. *There is a constant $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$,*

$$\left| -\log \mathbf{m}(\mathcal{N}_{\alpha,\beta}) - \min(\alpha, \mu(\beta)) \right| \leq c \log(\alpha + 2).$$

Since $\mu(\beta) > \alpha$ if and only if $\beta \geq \mathrm{BB}(\alpha)$, Theorem 5 says that

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}) = \begin{cases} 2^{-\mu(\beta) \pm O(\log(\alpha+2))} & \text{if } \beta < \mathrm{BB}(\alpha), \\ 2^{-\alpha \pm O(\log(\alpha+2))} & \text{otherwise.} \end{cases}$$

Therefore, for every computable bound $\beta \leq g(\alpha)$, $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ is only polynomially small in α (and Theorem 1 gives the more precise $\Theta(\tau(\alpha))$). As β grows beyond all computable bounds, $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ decreases as $2^{-\mu(\beta)}$, and it settles at the order $2^{-\alpha}$, familiar from the results discussed in the introduction, when β reaches $\mathrm{BB}(\alpha)$. See the end of Sect. A.6 for the proof of Theorem 5. Another expression for $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ working for all β will be given as (29) below.

The conditional universal measure of nonstochastic objects

The next theorem estimates the universal measure of nonstochastic objects above a given complexity level. Requiring $K(\omega) > k$ in addition to nonstochasticity simply replaces α in the bounds by $\max(\alpha, k)$, with no condition on k whatsoever. Theorem 6 below therefore contains Theorem 1, which is its special case for $k := 0$.

Theorem 6. *There is a constant $c > 1$ such that, for all $\alpha \in \mathbb{N}_0$, all $\beta \leq 2^{2^\alpha}$, and all $k \in \mathbb{N}_0$,*

$$\tau(\max(\alpha, k))/c \leq \mathbf{m}\big(\mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) > k\}\big) \leq c\tau(\max(\alpha, k)). \tag{7}$$

Only the lower bound in (7) uses a condition on β. See Sect. A.3 for the proof of Theorem 6.

We can see from (7) that α and k do not interact, and the larger of the two wins outright. Suppose $k \geq \alpha$ (the case $k < \alpha$ is almost vacuous, since every (α, β)-nonstochastic object satisfies $K(\omega) \geq \alpha - O(1)$). Comparing (7) with $\sigma(k) = \Theta(\tau(k))$ (see (3)) shows that a significant part of the mass that $\mathbf{m}$ assigns to the objects of complexity exceeding k is assigned to nonstochastic ones. Theorem 1 says that nonstochastic objects are generated with a probability that decays very slowly; Theorem 6 adds that, above the threshold α, nonstochastic objects are not merely plentiful: above any level $k \geq \alpha$ they carry a constant fraction of the universal measure. Formally, Theorem 6 implies that there is a constant $c > 0$ such that, for all $k \in \mathbb{N}_0$, all $\alpha \leq 2^{2^k}$, and all $\beta \leq 2^{2^\alpha}$,

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta} \mid K > k) \geq c, \tag{8}$$

where, as usual, $\mathbf{m}(A \mid B) := \mathbf{m}(A \cap B)/\mathbf{m}(B)$ and $K > k$ stands for $\{\omega \mid K(\omega) > k\}$.

3 Probability measures as models

A popular alternative to Kolmogorov's definition of stochastic objects uses probability measures in place of finite sets as models (see, e.g., [15] or [9, Sect. V]). For binary strings of length n, the two classes of models lead to the same notion of stochasticity to within $O(\log n)$ ([15, Theorem 1], [9, Proposition I.2]). In our structure-free setting there is no n, and changing the class of models has a dramatic effect on the universal measure of the nonstochastic objects. First we consider the simplest probability measures: a probability measure Q on $\mathcal{X}$ is *elementary* if it has a finite support and takes only rational values, as in [15]. (Much of this discussion also applies to general computable

probability measures as models, as in [9, Sect. V]; they are briefly covered at the end of this section.) Elementary probability measures are finite objects, and so $K(Q)$ and $K(\omega \mid Q)$ are defined.

An object $\omega \in \mathcal{X}$ is (α, β)-*quasi-stochastic*, for $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty]$, if there exists an elementary probability measure Q such that

- $K(Q) \le \alpha$;
- $Q(\{\omega\}) > 0$ and $K(\omega \mid Q) \ge -\log Q(\{\omega\}) - \beta$.

Let $\mathcal{N}^{\mathrm{q}}_{\alpha,\beta} \subseteq \mathcal{X}$ be the set of all objects that are not (α, β)-quasi-stochastic (or (α, β)-*non-quasi-stochastic*). Regarding a finite set Ω as the uniform probability measure on Ω, we can see that $\mathcal{N}^{\mathrm{q}}_{\alpha+c,\beta+c} \subseteq \mathcal{N}_{\alpha,\beta}$ for some constant c. The following version of Epstein and Levin's results (see below) shows that, in terms of the universal measure, $\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$ is much smaller than $\mathcal{N}_{\alpha,\beta}$.

Theorem 7. *There exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$,*

$$2^{-\alpha}/c \le \mathbf{m}(\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}) \le c\left(2^{-\alpha+K(\alpha)} + 2^{-2^{\beta/2-c}}\right).$$

There is a constant c' such that $2^{-2^{\beta/2-c}} \le 2^{-\alpha}$ whenever $\beta \ge 2\log(\alpha+1) + c'$. In this range of β,

$$\mathbf{m}(\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}) = O(2^{-\alpha+K(\alpha)}), \tag{9}$$

and so the contrast between the two classes of models is stark:

- with probability measures as models, non-quasi-stochastic objects are generated by the universal machine with an exponentially small probability; the term $2^{-\alpha}$, familiar from the results discussed in the introduction, reappears;
- by Theorem 1, $\mathbf{m}(\mathcal{N}_{\alpha,\beta}) = \Theta(\tau(\alpha))$ (for $\beta \le 2^{2^\alpha}$) is only polynomially small in α, and so nonstochastic objects are generated by the universal machine with a probability that tends to 0 very slowly.

See Sects. A.4 and A.5 for the proofs of the results in this section. The asymptotic relation (9) is a version of Epstein and Levin's results first posted in [6] (see [5, Lemma 10] and [10, Proposition 1] for published proofs). As mentioned in Sect. 1, those results might go back to Levin's idea explained in one of his notes to Kolmogorov likely written in 1970 [3, Sect. 1]. One of their differences from Theorem 7 is that Epstein and Levin merge the two standard parameters, α and β, into one and concentrate on a single number as their measure of stochasticity for an object ω, such as

$$\Lambda(\omega) := \min_Q \big(K(Q) + 3\log\max(d_{\mathrm{e}}(\omega \mid Q), 1)\big)$$

[5, Definition 1], where Q ranges over the elementary probability measures and d_{e} is defined in (11) below. Bauwens's report [2] contains all components needed for a slightly weaker version of the upper bound in Theorem 7 but does not state its results in terms of quasi-stochasticity.

The proof of Theorem 7 will show that we can improve its lower bound: there exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$,

$$\mathbf{m}(\mathcal{N}^{\mathrm{q}}_{\alpha,\beta} \cap \{\omega \mid K(\omega) \le \alpha + c\}) \ge 2^{-\alpha}/c.$$

Therefore, an $\mathbf{m}$-mass of $2^{-\alpha}/c$ (the lower bound in Theorem 7) is concentrated already on the $\omega \in \mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$ whose complexity $K(\omega)$ falls in the narrow range $[\alpha - c, \alpha + c]$. The proof will also show that it is concentrated on a single object.

Another modification of the definition of (α,β)-quasi-stochastic objects is to replace the inequality $K(\omega \mid Q) \geq -\log Q(\{\omega\}) - \beta$ by the Kolmogorov-style unconditional inequality $K(\omega) \geq -\log Q(\{\omega\}) - \beta$. Let $\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}$ be the set of all (α,β)-non-quasi-stochastic objects under the modified definition; we have $\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta+c} \subseteq \mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$ for some $c > 0$. This definition does not appear particularly natural to me, but it leads to the following elegant version of Theorem 7, whose upper bound is optimal (which implies the optimality in α of the upper bound in Theorem 7 itself).

Theorem 8. *There exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0,\infty]$,*

$$2^{-\alpha-c} \leq \mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}) \leq 2^{-\alpha+K(\alpha)+c}. \tag{10}$$

Moreover, if $f : \mathbb{N}_0 \to \mathbb{N}_0$ is an upper semicomputable function such that $\sum_n 2^{-f(n)} = \infty$, then, for infinitely many α, $\mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}) > 2^{-\alpha+f(\alpha)}$ for all $\beta \in [0,\infty]$.

As compared with Theorem 7, the bounds in (10) do not depend on β, and the upper bound is sharper in that the second addend in it disappears. The second part of Theorem 8, which follows from a result by Gács [7] (see also [1, Theorem 2]), shows that this upper bound cannot be improved. Indeed, every upper semicomputable $f : \mathbb{N}_0 \to \mathbb{N}_0$ with $\sum_n 2^{-f(n)} < \infty$ satisfies $f \geq K - O(1)$, and so, for any upper semicomputable $f : \mathbb{N}_0 \to \mathbb{N}_0$,

$$\mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}) = O\bigl(2^{-\alpha+f(\alpha)}\bigr) \iff \sum_n 2^{-f(n)} < \infty.$$

This shows that the dependence of $\mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta})$ on α is irregular: by (10), $\mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}) \leq 2^{-\alpha+O(\log\log\alpha)}$ when α is a power of 2, whereas $\mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}) > \alpha 2^{-\alpha}$ for infinitely many α (take $f(n) := \lceil \log(n+1) \rceil$).

General computable probability measures

Here we will use arbitrary computable probability measures on $\mathcal{X}$, as in [9, Sect. V], in place of elementary probability measures (used starting from [15]) as the models.

Let us fix a Gödel numbering (see, e.g., [17, Chap. 4]) of all algorithms that accept pairs in $\mathcal{X} \times \mathbb{N}$ as input and output rational numbers. We say that $\pi \in \mathbb{N}$ is a number of a probability measure P on $\mathcal{X}$ if it is a number of an algorithm that, given $\omega \in \mathcal{X}$ and $n \in \mathbb{N}$, outputs a rational number within 2^{-n} of $P(\{\omega\})$. An object $\omega \in \mathcal{X}$ is (α,β)-*Quasi-stochastic* (with a capital Q), where $\alpha \in \mathbb{N}_0$ and $\beta \in [0,\infty)$, if there exists a probability measure P on $\mathcal{X}$ with a number π such that

- $K(\pi) \leq \alpha$;
- $K(\omega \mid \pi) \geq -\log P(\{\omega\}) - \beta$

(with $-\log 0 := \infty$, as usual). Let $\mathcal{N}^{\mathrm{Q}}_{\alpha,\beta} \subseteq \mathcal{X}$ be the set of all objects that are not (α,β)-Quasi-stochastic. The definition does not depend on the chosen Gödel numbering, apart from additive constants in α and β, because any two Gödel numberings translate into each other computably.

For some constant $c > 0$, we always have $\mathcal{N}^{\mathrm{Q}}_{\alpha+c,\beta+c} \subseteq \mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$, since a number of an elementary probability measure Q can be computed from Q. Therefore, the upper bounds of Theorems 7 and 8 also hold for $\mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}$ and $\mathcal{N}^{\mathrm{Qu}}_{\alpha,\beta}$, respectively (the definition of the latter is obvious); it is not clear how to improve these upper bounds for general computable measures, since the Epstein–Levin methods produce elementary measures anyway. The lower bounds, however, have to be established anew. Remember that μ is the inverse busy beaver function (5) of Sect. 2. We state the analogue of the lower bound in Theorem 7; naturally, this bound tends to 0 as $\beta \to \infty$, unlike the lower bound of Theorem 7 (there exist computable probability measures with the whole of $\mathcal{X}$ as their support).

Theorem 9. *There exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$,*

$$\mathbf{m}(\mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}) \geq 2^{-\alpha-\mu(\beta)-c}.$$

In the proof of Theorem 9 in Sect. A.5 we will construct $\omega^* \in \mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}$ such that $K(\omega^*) \leq \alpha + \mu(\beta) + c$, so that the lower bound in Theorem 9 is attained on one object. The proof will also show that Theorem 9 can be strengthened by replacing $\mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}$ with $\mathcal{N}^{\mathrm{Qu}}_{\alpha,\beta}$, obtained by replacing the conditional complexity $K(\omega \mid \pi)$ in the second condition in the definition of (α, β)-Quasi-stochasticity with the unconditional complexity $K(\omega)$.

4 A generalization of Theorem 1

In this section we will see stronger and simpler versions of our first key result, Theorem 1. But before discussing new results, let us sort out our notation.

The most basic set related to stochasticity in this paper is $\mathcal{N}_{\alpha,\beta}$ introduced in Sect. 2. Its modification $\mathcal{N}^{C}_{\alpha,\beta}$, obtained by replacing K with C in both the α-part and the β-part of the definition of $\mathcal{N}_{\alpha,\beta}$, was also discussed in Sect. 2. Kolmogorov was interested in the set $\mathcal{N}^{C\mathrm{u}}_{\alpha,\beta}$ resulting from replacing the conditional complexity $C(\omega \mid \Omega)$ in the β-part of the definition of $\mathcal{N}^{C}_{\alpha,\beta}$ with the unconditional complexity $C(\omega)$; this will be further discussed in Appendix B. In Sect. 3, we discussed modifications $\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$ and $\mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}$ obtained by replacing finite sets as models with probability measures, and we also used $\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta}$ in Theorem 8 and $\mathcal{N}^{\mathrm{Qu}}_{\alpha,\beta}$ right after Theorem 9. In this section we will be interested in various other modifications of $\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$, so that our models here are elementary probability measures, and we will need a richer and more flexible system of notation; the new system, however, will not cover the sets $\mathcal{N}^{\mathrm{Q}}$.

We are only interested in two versions of the α-part, namely using C or K. As for the β-part, there is a wider range of possibilities. Let us define a *deficiency* to be a lower semicomputable function d mapping each pair (ω, Q), where $\omega \in \mathcal{X}$ and Q is an elementary probability measure, to $d(\omega \mid Q) \in (-\infty, \infty]$. Then we say that an object ω is (α, β)-*quasi-stochastic w.r.t.* d, where $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$, if there exists an elementary probability measure Q such that $K(Q) \leq \alpha$ and $d(\omega \mid Q) \leq \beta$. Define $\mathcal{N}_{\alpha,\beta}(d)$ to be the set of all objects that are not (α, β)-quasi-stochastic w.r.t. d. A natural modification is to define $\mathcal{N}^{C}_{\alpha,\beta}(d)$ by replacing K with C in the definition of $\mathcal{N}_{\alpha,\beta}(d)$.

We will be interested in several specific deficiencies and classes of deficiencies. The basic set $\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$ is identical to $\mathcal{N}_{\alpha,\beta}(d_{\mathrm{e}})$ with

$$d_{\mathrm{e}}(\omega \mid Q) := -\log Q(\{\omega\}) - K(\omega \mid Q) \tag{11}$$

(the origin of the lower index "e" is the connection with e-values, to be discussed in Sect. 5 below). For an upper semicomputable function $F : \mathcal{X} \to \mathbb{N}_0$, set

$$d_F(\omega \mid Q) := -\log Q(\{\omega\}) - F(\omega), \tag{12}$$

with $-\log 0 := \infty$. In this notation, $\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta} = \mathcal{N}_{\alpha,\beta}(d_K)$.

To embed our basic set $\mathcal{N}_{\alpha,\beta}$ used in Theorem 1 in the new notation, define

$$d^{\mathrm{set}}(\omega \mid Q) := \begin{cases} d(\omega \mid Q) & \text{if } Q = U_\Omega \text{ for a finite set } \Omega \subseteq \mathcal{X} \\ \infty & \text{otherwise} \end{cases}$$

for a given deficiency d, where U_Ω is the uniform probability measure on Ω. Then $\mathcal{N}_{\alpha,\beta}$ and $\mathcal{N}_{\alpha,\beta}(d_{\mathrm{e}}^{\mathrm{set}})$ coincide to within additive constants: for some constant $c > 0$ and for all α and β,

$$\mathcal{N}_{\alpha+c,\beta+c} \subseteq \mathcal{N}_{\alpha,\beta}(d_{\mathrm{e}}^{\mathrm{set}}) \quad \text{and} \quad \mathcal{N}_{\alpha+c,\beta+c}(d_{\mathrm{e}}^{\mathrm{set}}) \subseteq \mathcal{N}_{\alpha,\beta}.$$

Kolmogorov's original set $\mathcal{N}_{\alpha,\beta}^{C\mathrm{u}}$ embeds into the new notation as $\mathcal{N}_{\alpha,\beta}^{C}(d_C^{\mathrm{set}})$. Indeed, $\mathcal{N}_{\alpha,\beta}^{C\mathrm{u}}$ and $\mathcal{N}_{\alpha,\beta}^{C}(d_C^{\mathrm{set}})$ coincide to within additive constants: for some constant $c > 0$ and for all α and β,

$$\mathcal{N}_{\alpha+c,\beta+c}^{C\mathrm{u}} \subseteq \mathcal{N}_{\alpha,\beta}^{C}(d_C^{\mathrm{set}}) \quad \text{and} \quad \mathcal{N}_{\alpha+c,\beta+c}^{C}(d_C^{\mathrm{set}}) \subseteq \mathcal{N}_{\alpha,\beta}^{C\mathrm{u}}.$$

Notice that the new notation using "(d)" is more flexible than the old notation, since, e.g., now we can use different complexities in the α-part and β-part, such as using C in the α-part and K in the β-part.

The following theorem says that non-quasi-stochastic objects are frequent whenever a deficiency penalizes models at least as heavily as d_F, to within multiplicative and additive constants, for some F making, unlike K, the series $\sum_\omega 2^{-F(\omega)}$ divergent.

Theorem 10. *Let d be a deficiency and $F : \mathcal{X} \to \mathbb{N}_0$ be an upper semicomputable function satisfying $\sum_\omega 2^{-F(\omega)} = \infty$. Suppose there is a constant $c > 0$ such that*

$$d_F(\omega \mid Q) \le c d(\omega \mid Q) + c \tag{13}$$

for all ω and all elementary probability measures Q. There exists $c' > 0$ such that $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d)) \ge \tau(\max(\alpha,\beta))/c'$ for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0,\infty)$.

The non-trivial part of Theorem 1 (the lower bound) will follow immediately from Theorem 10 since $d_C \le 2d_{\mathrm{e}}^{\mathrm{set}} + O(1)$. Theorem 10 obviously remains true for $\mathbf{m}(\mathcal{N}_{\alpha,\beta}^{C}(d))$ in place of $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d))$.

For the deficiencies (12), Theorems 8 and 10 lead to the following dichotomy.

Theorem 11. *Let $F : \mathcal{X} \to \mathbb{N}_0$ be an upper semicomputable function.*

- *If $\sum_\omega 2^{-F(\omega)} < \infty$, there is a constant $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and all $\beta \in [0,\infty)$,*

$$2^{-\alpha-c} \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) \le 2^{-\alpha+K(\alpha)+c}. \tag{14}$$

  (Only the upper bound depends on the condition on F.)

- *If $\sum_\omega 2^{-F(\omega)} = \infty$, there is a constant $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and all $\beta \in [0,\infty)$,*

$$\tau(\max(\alpha,\beta))/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) \le c\tau(\alpha), \tag{15}$$

  and

$$|-\log \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) - \min(\alpha, \mu(\beta))| \le c \log(\alpha+2). \tag{16}$$

Prefix complexity K and plain complexity C represent different sides of the dichotomy in Theorem 11. An example demonstrating an anomalous feature of d_C (favoured by Kolmogorov in the context of finite sets as models) as a measure of randomness was given by Levin [3, Sect. 4.2] in his undated letter to Kolmogorov (written between August 1970 and January 1971): define a probability measure on the binary strings by

$$Q(\{x\}) = \begin{cases} 2^{-|x|-100} & \text{if } |x| \le 2^{100} \\ 0 & \text{otherwise;} \end{cases}$$

then $-\log Q(\{x\}) - C(x) > 100$ for all x. (Formally, the binary strings should be represented by $\{0,1\}^* \setminus \{\Box\}$, as on $\{0,1\}^*$ the total mass of Q would be $1 + 2^{-100}$, and $-\log Q(\{x\}) - C(x) > 100$ should be understood as $-\log Q(\{x\}) - C(x) > 100 - c$ for some constant c, but this does not matter as 100 can be replaced by an arbitrarily large number.)

The function C can be increased considerably without leaving the C-side of the dichotomy, as demonstrated by the following lemma.

Lemma 12. *For a computable increasing function $h : \mathbb{N}_0 \to \mathbb{N}_0$,*

$$\sum_{\omega} 2^{-C(\omega)-h(C(\omega))} = \infty \iff \sum_{n} 2^{-h(n)} = \infty.$$

Proof. If $\sum_n 2^{-h(n)} < \infty$, then $K \le h + O(1)$ on $\mathbb{N}_0$, and so $C + h(C) \ge C + K(C) - O(1) \ge K - O(1)$, which implies $\sum_\omega 2^{-C(\omega)-h(C(\omega))} < \infty$. On the other hand, if $\sum_n 2^{-h(n)} = \infty$, considering, for each n, the 2^n objects in $[2^{n+1}] \setminus [2^n]$ gives, for some $c > 0$,

$$\sum_{\omega} 2^{-C(\omega)-h(C(\omega))} \ge \sum_{n=0}^{\infty} 2^n 2^{-n-c-h(n+c)} = 2^{-c} \sum_{m=c}^{\infty} 2^{-h(m)} = \infty. \qquad \Box$$

For example, d_F leads to frequent non-quasi-stochastic objects for $F := C + \lceil \log(C+1) + \log\log(C+2) \rceil$ but to rare ones for $F := C + \lceil \log(C+1) + 2\log\log(C+2) \rceil$.

5 Omnibus theorems

In this section we complement the deficiencies d_{e} and d_F introduced in the previous section by the important deficiency proposed, in slightly different forms, by Martin-Löf [11]. Let us say that a function t mapping each pair (ω, Q), where $\omega \in \mathcal{X}$ and Q is an elementary probability measure, to $\mathbb{N}_0 \cup \{\infty\}$ is a *randomness p-test* if it is lower semicomputable and satisfies

$$Q(\{\omega \mid t(\omega, Q) \ge n\}) \le 2^{-n}$$

for all Q and all $n \in \mathbb{N}$. Let d_{p} be a largest, to within an additive constant, randomness p-test. We fix d_{p} and refer to $d_{\mathrm{p}}(\omega \mid Q)$ as the *randomness p-deficiency* of ω under Q. For the uniform probability measures U_Ω on finite sets Ω,

$$d_{\mathrm{p}}(\omega \mid U_\Omega) = \log|\Omega| - C(\omega \mid \Omega) + O(1) \tag{17}$$

for $\omega \in \Omega$ (this is Martin-Löf's expression of the conditional version of Kolmogorov's deficiency in terms of tests [11, Sect. II, Theorem on p. 607]; see also [19, Appendix 1, Proposition 11]).

The randomness p-deficiency $d_{\mathrm{p}}(\cdot \mid Q)$ is a p-variable for Q on the logarithmic scale, meaning that $2^{-d_{\mathrm{p}}(\cdot|Q)}$ is a p-variable for Q, as defined in, e.g., [13, Definition 1.2(ii)]. Similarly, the function $d_{\mathrm{e}}(\cdot \mid Q)$ defined by (11) is an e-variable for Q on the logarithmic scale, meaning that $2^{d_{\mathrm{e}}(\cdot|Q)}$ is an e-variable for Q, as defined in, e.g., [13, Definition 1.2(i)]. Both d_{p} and d_{e} are the largest, to within an additive constant, lower semicomputable functions in these classes (of the p- and e-variables, respectively, on the logarithmic scale). Therefore, it is natural to call $d_{\mathrm{e}}(\omega \mid Q)$ the *randomness e-deficiency* of ω under Q. The functions d_{p} and d_{e} can be interpreted as universal versions of p-values and e-values, respectively.

Now we can state our first omnibus theorem covering eight cases; it will be proved in Sect. A.7.

Theorem 13. *Let $F : \mathcal{X} \to \mathbb{N}_0$ be an upper semicomputable function such that $\sum_\omega 2^{-F(\omega)} = \infty$ (e.g., $F = C$). There exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$,*

$$2^{-\alpha}/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_{\mathrm{e}})) \le c\left(2^{-\alpha+K(\alpha)} + 2^{-2^{\beta/2-c}}\right), \tag{18}$$

$$2^{-\alpha}/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_{\mathrm{p}})) \le c\left(2^{-\alpha+K(\alpha)} + 2^{-2^{\beta/2-c}}\right), \tag{19}$$

$$2^{-\alpha}/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_K)) \le c2^{-\alpha+K(\alpha)}, \tag{20}$$

$$\left(2^{-\alpha} + \tau(\max(\alpha,\beta))\right)/c \le \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) \le c\left(2^{-\alpha+K(\alpha)} + \tau(\max(\alpha,\beta))\right), \tag{21}$$

and

$$2^{-\alpha-K(\alpha)}/c \le \mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d_{\mathrm{e}})) \le c\left(2^{-\alpha} + 2^{-2^{\beta/2-c}}\right), \tag{22}$$

$$2^{-\alpha-K(\alpha)}/c \le \mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d_{\mathrm{p}})) \le c\left(2^{-\alpha} + 2^{-2^{\beta/2-c}}\right), \tag{23}$$

$$2^{-\alpha-K(\alpha)}/c \le \mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d_K)) \le c2^{-\alpha}, \tag{24}$$

$$\left(2^{-\alpha-K(\alpha)} + \tau(\max(\alpha,\beta))\right)/c \le \mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d_F)) \le c\left(2^{-\alpha} + \tau(\max(\alpha,\beta))\right). \tag{25}$$

The eight double inequalities in Theorem 13 have an interesting structure. The bounds for d_{e} and d_{p} are identical, for both conditions $K(Q) \le \alpha$ and $C(Q) \le \alpha$, despite e-values and p-values being regarded as very different in statistics. Replacing the condition $K(Q) \le \alpha$ with $C(Q) \le \alpha$ moves the factor $2^{K(\alpha)}$ from the upper bound to the lower bound. For d_{e}, d_{p}, and d_K, non-quasi-stochastic objects are rare: their universal measure is $2^{-\alpha+O(\log(\alpha+2))}$ (for d_{e} and d_{p}, as soon as $\beta \ge 2\log(\alpha+1) + c'$). For d_F, they are frequent, in agreement with Theorem 11. The first sandwich, (18), is Theorem 7, and (20) is the first part of Theorem 8.

The terms $2^{-\alpha}$ and $2^{-\alpha\pm K(\alpha)}$ in (21) and (25) matter only for huge β. For a fixed computable function g and all $\beta \le g(\alpha)$, they are dominated by the terms $\tau(\max(\alpha,\beta))$, and both measures in (21) and (25) are $\Theta(\tau(\alpha))$ (by Lemma 3), as in Theorem 1. However, $\tau(\max(\alpha,\beta)) \to 0$ as $\beta \to \infty$, while the measures in (21) and (25) never fall below $2^{-\alpha}/c$ and $2^{-\alpha-K(\alpha)}/c$, respectively. In (16), these terms correspond to the α in $\min(\alpha, \mu(\beta))$.

The upper bounds in (21) and (25) hold for any upper semicomputable $F : \mathcal{X} \to \mathbb{N}_0$ (only $F \ge 0$ is used); the condition $\sum_\omega 2^{-F(\omega)} = \infty$ is only needed for the terms $\tau(\max(\alpha,\beta))$ in the lower bounds. The sandwich (21) implies both (15) and (16), but we will give simple independent derivations for the latter two in Sect. A.6.

Theorem 13 is about non-quasi-stochastic objects. The situation for nonstochastic objects is much simpler, as the following result (proved in Sect. A.8) shows.

Theorem 14. *Let F be as in Theorem 13. There exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$, all $\beta \in [0, \infty)$, and each $d \in \{d_{\mathrm{e}}, d_{\mathrm{p}}, d_K, d_F\}$,*

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})) \ge \left(2^{-\alpha} + \tau(\max(\alpha,\beta))\right)/c, \tag{26}$$

$$\mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}})) \ge \left(2^{-\alpha-K(\alpha)} + \tau(\max(\alpha,\beta))\right)/c, \tag{27}$$

and both measures are at most

$$c\left(2^{-\alpha+c\log(\alpha+2)} + \tau(\max(\alpha,\beta))\right). \tag{28}$$

In all eight cases covered by Theorem 14, (α,β)-nonstochastic objects are frequent under the universal measure provided $\beta \le g(\alpha)$ for a computable g: their universal measure is $\Theta(\tau(\alpha))$. Moreover, the theorem (namely, (26) and (28) with $d := d_{\mathrm{e}}$) shows that our first key result (2) can be strengthened to

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}) = \Theta\left(\tau(\max(\alpha,\beta)) + 2^{-\alpha+O(\log(\alpha+2))}\right), \tag{29}$$

the second addend on the right-hand side kicking in for extremely large β. This expression for $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ complements Theorem 5, which was given in terms of μ rather than τ.

Statements in terms of the conditional universal measure

In Theorem 1 and further results of this kind, such as (21), (25), and (26)–(27), we interpreted a set such as $\mathcal{N}_{\alpha,\beta}$ having an $\mathbf{m}$-share of $\tau(\alpha)$ as being frequent. However, $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ still tends to 0 as $\alpha \to \infty$, albeit slowly. In this subsection these results will be restated in terms of a constant proportion of nonstochastic or non-quasi-stochastic objects.

Theorem 15. *Let F be as in Theorem 13. There exists $c > 0$ such that, for all $\alpha, k \in \mathbb{N}_0$ and $\beta \in [0,\infty)$, each of the bounds in Theorems 13 and 14, of the form $\mathbf{m}(\mathcal{N}) \ge L/c$ or $\mathbf{m}(\mathcal{N}) \le cU$, remains true when $\mathbf{m}(\mathcal{N})$ is replaced by $\mathbf{m}(\mathcal{N} \cap B)$, where B is either $\{\omega \mid K(\omega) > k\}$ or $\{\omega \mid C(\omega) > k\}$, and L and U are replaced by $\min(L, \tau(k))$ and $\min(U, \tau(k))$, respectively.*

See Sect. A.9 for the proof. In view of Lemma 3 and (3), Theorem 15 immediately implies the following corollary.

Corollary 16. *Let F be as in Theorem 13. There exists $c > 0$ such that, for all $k \in \mathbb{N}_0$ and all $\alpha \in \mathbb{N}_0, \beta \in [0,\infty)$ satisfying $\max(\alpha,\beta) \le 2^{2^k}$, we have $\mathbf{m}(A \mid B) > c$ whenever*

- *A is one of the sets $\mathcal{N}_{\alpha,\beta}(d_F)$, $\mathcal{N}^C_{\alpha,\beta}(d_F)$, $\mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})$, or $\mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}})$, with $d \in \{d_{\mathrm{e}}, d_{\mathrm{p}}, d_K, d_F\}$, and*
- *B is either $\{\omega \mid K(\omega) > k\}$ or $\{\omega \mid C(\omega) > k\}$.*

As before, the tower 2^{2^k} can be replaced by $g(k)$ for any computable increasing g; for a higher tower, (8) becomes essentially a special case of this corollary. The corollary shows that the non-stochastic objects, as well as non-quasi-stochastic objects defined in terms of d_F, make up at least a constant $\mathbf{m}$-fraction of all objects of high complexity. The statement about d_F is the main one for the proofs in this paper, since we reduce the cases of d^{set} to d_F via Theorem 10.

6 Conclusion

Let me summarize the main features of this paper's structure-free picture. The universal measure of (α,β)-nonstochastic objects hardly depends on β (unless β is huge, in a busy-beaver sense, as compared with α) and tends to 0 extremely slowly as $\alpha \to \infty$, more slowly than any decreasing computable function that tends to 0. Above any complexity level $k \ge \alpha$ a significant part of $\mathbf{m}$ is allocated to (α,β)-nonstochastic objects:

$$\mathbf{m}\big(\mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) > k\}\big) = \Theta\big(\mathbf{m}(\{\omega \mid K(\omega) > k\})\big)$$

(provided β is computably bounded in α). In sharp contrast with (α,β)-nonstochastic objects, the universal measure of (α,β)-non-quasi-stochastic objects is exponentially small in α as $\alpha \to \infty$ provided $\beta \ge 2\log\alpha$, both in the case where the quality of models is measured by universal e-values and in the case where it is measured by universal p-values.

A Proofs

In this appendix we will prove all the theorems stated in the main part of this paper. The basic space $\mathcal{X}$ is interpreted as $\mathbb{N}$.

In the proofs, $c_1, c_2, \dots$ will stand for various positive constants; numbering them facilitates checking the absence of vicious circles in their definitions. Each constant is absolute or depends only on previously defined constants, except for a few explicitly chosen "large enough" later in the same proof, after constants that do not depend on them.

Let me first describe the main techniques used in the proofs in this appendix and in the proof of Theorem 25 in Appendix B. Proofs of the upper bounds often rely on the standard techniques of halting cuts, either cuts of Chaitin's number $\mathbf{\Omega}$ as used, e.g., by Epstein and Levin [5, Lemma 10], or cuts of "bounded complexity lists" (see, e.g., [9, Sect. III.E] and [18, Sect. 4]). These standard techniques are used in the proofs of the upper bounds of Theorems 7–8, 11, 13–15, and 25 (sometimes indirectly).

The main new technique introduced in this paper (with an important contribution from AI tools) is used in proofs of lower bounds of numerous theorems (including, directly or indirectly, Theorems 1, 5–6, 10–11, and 13–15). It develops Gács et al.'s observation that the complexity of stochastic objects is simple [9, Lemmas III.8 and III.10]: $K(K(\omega))$ cannot be very large as compared with α and β for (α, β)-stochastic objects. Combined with the fact that the universal measure often generates ω with a large $K(K(\omega))$, this implies that nonstochastic objects are abundant. Early versions of this paper (such as v1 and v2 of the arXiv report) relied literally on this observation. This mutated into using the observation that, for stochastic objects ω, $K(\omega) \approx C(\omega)$, while the universal measure often generates ω with very different $K(\omega)$ and $C(\omega)$. Because of Solovay's result

$$K(\omega) - C(\omega) = K(K(\omega)) + O(K(K(K(\omega))))$$

(see, e.g., [16, Sect. 4.7.4]), there is very little difference between the two observations. It turns out that in the second observation C can be replaced by any upper semicomputable function F such that $\sum_\omega 2^{-F(\omega)} = \infty$ (see (40) in the proof of Theorem 10 and Lemma 18 below). Now the properties $K(K(\omega)) \approx 0$ and $K(\omega) \approx F(\omega)$ of an object ω become genuinely different.

Another new technique used for several lower bounds involves a simple modification of Shen's technique of the "first uncovered object" [15, Theorem 2(b)]. Shen's technique was also used in early versions of this paper (including v1 and v2 of the arXiv report), but its simplification (discovered with the help of AI tools) has strengthened several results. It is used in proving the lower bounds in Theorems 5, 7–9, 11, 13–15, and 25.

There are also several idiosyncratic new techniques, each used in just one proof.

A.1 The tail of the universal measure

This section proves key results about the function τ, which plays a prominent role in several bounds, and related functions. Some of the proofs employ the great stability of τ; it hardly changes over huge regions of its domain. We start with the trivial observation that the composition of a computable Markov kernel with the universal measure is dominated by the universal measure.

Lemma 17. *Let $(\kappa_n)_{n\in\mathbb{N}}$ be probability measures on $\mathbb{N}$ such that the mapping $(n, i) \in \mathbb{N}^2 \mapsto \kappa_n(\{i\})$ is computable. There is $c_\kappa > 0$ such that, for all $A \subseteq \mathbb{N}$,*

$$\mathbf{m}(A) \geq c_\kappa \sum_{n\in\mathbb{N}} \mathbf{m}(n)\kappa_n(A).$$

Proof. The function $i \mapsto \sum_n \mathbf{m}(n)\kappa_n(\{i\})$ is a lower semicomputable semimeasure, and the universality of $\mathbf{m}$ gives $\mathbf{m}(i) \geq c_\kappa \sum_n \mathbf{m}(n)\kappa_n(\{i\})$ for all i. It remains to sum over $i \in A$. □

Our first application of Lemma 17 is to demonstrate the great stability of τ asserted in Lemma 3.

Proof of Lemma 3. Applying Lemma 17 to κ_n set to $\delta_{g(n)+1}$, the point mass at $g(n)+1$, we have

$$\tau(g(t)) = \mathbf{m}(\{i \mid i > g(t)\}) \geq c_\kappa \sum_{n\in\mathbb{N}} \mathbf{m}(n) 1_{\{g(n)\geq g(t)\}}$$
$$\geq c_\kappa \sum_{n\in\mathbb{N}} \mathbf{m}(n) 1_{\{n\geq t\}} = c_\kappa \sum_{n\geq t} \mathbf{m}(n) \geq c_\kappa \tau(t). \quad \square$$

Let us check the relation between τ and σ stated in Sect. 2 directly (the lower bound on σ will also follow from the next lemma, but the direct proof is easier).

Proof of Lemma 4. The lower bound in (3) follows from Lemma 17 applied to $A := \{n \mid K(n) > t\}$ and the uniform probability measures κ_k on $[2^k]$. At most 2^t numbers n have $K(n) \leq t$, so for $k \geq t+1$ we have $\kappa_k(A) \geq 1/2$; therefore,

$$\sigma(t) = \mathbf{m}(\{n \mid K(n) > t\}) = \mathbf{m}(A) \geq c_1 \sum_{k=t+1}^{\infty} \frac{\mathbf{m}(k)}{2} = \frac{c_1}{2}\tau(t).$$

The upper bound follows from

$$\sigma(t) = \mathbf{m}(\{n \mid K(n) > t\}) \leq \mathbf{m}(\{n \mid n > t - c_2\}) = \tau(t - c_2) \leq c_3\tau(t),$$

where we have used Lemma 3; this chain assumes $t \geq c_2$ (smaller t are covered by $\sigma \leq 1$ and $\tau(t) > 0$). □

The following lemma (which will be one of our key technical tools) generalizes the lower bound in (3).

Lemma 18. *Let $F : \mathcal{X} \to \mathbb{N}_0$ be an upper semicomputable function such that $\sum_\omega 2^{-F(\omega)} = \infty$. There exists a constant $c_4 > 0$ such that, for all $t \in \mathbb{N}_0$,*

$$\mathbf{m}(\{\omega \mid K(\omega) - F(\omega) > t\}) \geq c_4\tau(t). \tag{30}$$

For $F := 0$, this is the lower bound in (3). On the other hand, if $\sum_\omega 2^{-F(\omega)} < \infty$, $K - F$ is bounded above and so the left-hand side of (30) vanishes from some t on.

Proof. Let F_s, $s \in \mathbb{N}_0$, be a computable approximation to F from above (so that $F_s(\omega) \in \mathbb{N}_0 \cup \{\infty\}$ decreases to $F(\omega)$ as $s \to \infty$). For $k \in \mathbb{N}$, find the first pair (s, n) (in a fixed computable enumeration of $\mathbb{N}_0 \times \mathbb{N}$) such that $Z_k := \sum_{\omega\in[n]} 2^{-F_s(\omega)} \geq 2^k$ (such pairs exist since $\sum_\omega 2^{-F(\omega)} = \infty$), and let κ_k be the probability measure on $[n]$ defined by $\kappa_k(\{\omega\}) := 2^{-F_s(\omega)}/Z_k$. Since $\kappa_k(\{\omega\}) \leq 2^{-k-F(\omega)}$ for all ω, we have

$$\kappa_k\big(\{\omega \mid K(\omega) - F(\omega) \leq t\}\big) \leq \sum_{\omega: F(\omega)\geq K(\omega)-t} 2^{-k-F(\omega)} \leq \sum_\omega 2^{-k-K(\omega)+t} \leq 2^{t-k}.$$

Therefore, $\kappa_k(A) \geq 1/2$ for $A := \{\omega \mid K(\omega) - F(\omega) > t\}$ and all $k \geq t+1$, and Lemma 17 gives

$$\mathbf{m}(A) \geq \frac{c_\kappa}{2} \sum_{k=t+1}^{\infty} \mathbf{m}(k) = \frac{c_\kappa}{2}\tau(t). \quad \square$$

The next lemma spells out the easy part of (6).

Lemma 19. *For some constant $c_5 > 0$, $\tau \ge c_5 2^{-\mu}$.*

Proof. The inequality follows from

$$\tau(t) = \mathbf{m}(\{n \mid n > t\}) \ge \mathbf{m}(n^*) \ge c_5 2^{-K(n^*)} = c_5 2^{-\mu(t)},$$

where $n^* > t$ is any number that achieves $K(n^*) = \mu(t)$. □

A.2 Proof of Theorem 1

We will prove Theorem 1 in the form (2). This implies the other version of Theorem 1 given right after its statement. This was already mentioned in the main paper, but let me spell out all details. Let g be a computable function; we assume, without loss of generality, that g is increasing and $g(n) \ge n$ for all n in its domain. Lemma 3 gives, for $\beta \le g(\alpha)$,

$$\tau(\max(\alpha, \beta)) \ge \tau(g(\alpha)) \ge c_g \tau(\alpha).$$

It remains to combine this with (2). The main statement of the theorem uses $g(\alpha) := 2^{2^\alpha}$.

Let us start with the easy upper bound in (2). For some constant c_6,

$$K(\{\omega\}) \le K(\omega) + c_6. \tag{31}$$

Every $\omega \in \mathcal{N}_{\alpha,\beta}$ satisfies $K(\omega) > \alpha - c_6$ (otherwise $\Omega := \{\omega\}$ would have witnessed its $(\alpha, 0)$-stochasticity). Therefore, every $\omega \in \mathcal{N}_{\alpha,0}$ satisfies $2 \log \omega \ge \alpha - c_7$, which implies

$$\omega \ge 2^{\alpha/2 - c_7/2} \ge \alpha + 1$$

for all $\alpha \ge c_8$. This implies

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}) \le \mathbf{m}(\mathcal{N}_{\alpha,0}) \le \mathbf{m}(\{n \mid n > \alpha\}) = \tau(\alpha)$$

from $\alpha = c_8$ on. (This is sufficient since the remaining finitely many αs can be absorbed into the constant in (2).)

It remains to prove the lower bound. It is an easy corollary of Theorem 10, which will be proved in Sect. A.6 independently of this proof. For some constants c_9 and c_{10}, for all finite sets $\Omega \subseteq \mathcal{X}$ and all $\omega \in \Omega$,

$$\log |\Omega| - C(\omega) \le \log |\Omega| - C(\omega \mid \Omega) + c_9 \le 2\Big(\log |\Omega| - K(\omega \mid \Omega)\Big) + c_{10} \tag{32}$$

([19, Proposition 17]; the proof in [12, Proposition 1] is applicable); therefore,

$$d_C(\omega \mid Q) \le 2 d_{\mathrm{e}}^{\mathrm{set}}(\omega \mid Q) + c_{11},$$

and Theorem 10 gives the lower bound.

The version of the theorem with $\mathcal{N}^C_{\alpha,\beta}$ in place of $\mathcal{N}_{\alpha,\beta}$ is proved in the same way (and the second inequality in (32) is not needed).

A.3 Proof of Theorem 6

Set $A := \mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) > k\}$. Let us check that, analogously to (2),

$$\tau(\max(\alpha,\beta,k))/c_{12} \le \mathbf{m}(A) \le c_{12}\tau(\max(\alpha,k)). \tag{33}$$

Under the condition $\beta \le g(\alpha)$ for a computable increasing g, Lemma 3 then gives

$$\mathbf{m}(A) \ge \tau(\max(\alpha,\beta,k))/c_{12} \ge \tau(\max(g(\alpha),k))/c_{12} \ge c_{13}\tau(\max(\alpha,k))$$

for some $c_{13} > 0$ (again assuming, without loss of generality, $g(n) \ge n$ for all n).

We start with the upper bound in (33). Since τ is decreasing, $\tau(\max(\alpha,k)) = \min(\tau(\alpha),\tau(k))$, so it suffices to bound $\mathbf{m}(A)$ above by a multiple of each of $\tau(\alpha)$ and $\tau(k)$. From $A \subseteq \mathcal{N}_{\alpha,\beta}$ and the upper bound in (1) we get $\mathbf{m}(A) \le c\tau(\alpha)$; this part of Theorem 1 needs no condition on β, as $\mathcal{N}_{\alpha,\beta} \subseteq \mathcal{N}_{\alpha,0}$. And since $A \subseteq \{\omega \mid K(\omega) > k\}$, (3) gives $\mathbf{m}(A) \le \sigma(k) \le c_{14}\tau(k)$.

It remains to establish the lower bound in (33). Set $\alpha' := \max(\alpha, k + c_6)$ (see (31) for c_6). Since $\alpha' \ge \alpha$, we have $\mathcal{N}_{\alpha',\beta} \subseteq \mathcal{N}_{\alpha,\beta}$; and since every $\omega \in \mathcal{N}_{\alpha',\beta}$ satisfies $K(\omega) > \alpha' - c_6 \ge k$, we have $\mathcal{N}_{\alpha',\beta} \subseteq A$. Therefore, the lower bound in (2) gives

$$\mathbf{m}(A) \ge \mathbf{m}(\mathcal{N}_{\alpha',\beta}) \ge \tau(\max(\alpha',\beta))/c_{15} \ge \tau\big(\max(\alpha,\beta,k) + c_6\big)/c_{15},$$

and Lemma 3 gives $\mathbf{m}(A) \ge \tau(\max(\alpha,\beta,k))/c_{12}$ for a large enough c_{12} (which should also cover the upper bound in (33)).

A.4 Proof of Theorems 7 and 8

Before we start the proof, let me explain in greater detail the reason for the stark contrast between Theorem 1 (there are many nonstochastic objects) and Theorem 7 (there are few non-quasi-stochastic objects). A finite set Ω used as a model determines $\log|\Omega|$, and so it essentially determines the complexity $K(\omega)$ of its typical elements $\omega \in \Omega$. As Ω is simple, the complexity of stochastic objects is also simple; the objects whose complexity is complex are nonstochastic. A probability measure Q, on the other hand, does not have to determine $-\log Q(\{\omega\})$: it can spread its mass over objects of different complexities.

Proof of the upper bound of Theorem 7. In this proof we employ the technique of halting cuts. Let U be the optimal prefix-free decoder, so that $K(\omega) = \min\{|p| \mid U(p) = \omega\}$ for all ω. Let $\mathbf{\Omega}$ (Chaitin's number Ω) be its halting probability, i.e., the sum of the *weights* $2^{-|p|}$ over all p such that $U(p)$ halts. Fix a process of computing $U(p)$ for all $p \in \{0,1\}^*$, and let $\mathbf{\Omega}_t$, $t \in \mathbb{N}_0$, be its halting probability in t steps, i.e., the sum of $2^{-|p|}$ over all p such that $U(p)$ halts in at most t steps. For each $k \in \mathbb{N}_0$, define

$$T_k := \min\left\{t \mid \mathbf{\Omega}_t \ge \lfloor 2^k\mathbf{\Omega}\rfloor 2^{-k}\right\};$$

in words, T_k is the time of computing the k-bit truncation of $\mathbf{\Omega}$. Next $k(p)$, $p \in \{0,1\}^*$, is the *slice* of p, defined to be the k such that the number of steps in which $U(p)$ halts belongs to the interval $(T_k, T_{k+1}]$. (This is well defined since $T_0 = 0$ and $T_k \to \infty$ as $k \to \infty$.)

Next we define our models. For $k \in \mathbb{N}_0$, define a measure Q_k on $\mathcal{X}$ by setting $Q_k(\{\omega\})$ to 2^k times the total weight of p in slice k satisfying $U(p) = \omega$. This definition only ensures $\sum_\omega Q_k(\{\omega\}) \le 1$, so we increase the value $Q_k(\{\omega_0\})$ for a fixed $\omega_0 \in \mathcal{X}$ to make Q_k a probability measure. The values of Q_k are dyadic rationals, so Q_k are elementary probability measures.

To bound $\mathbf{m}(\mathcal{N}^{\mathrm{q}}_{\alpha,\beta})$ from above, let us consider $\omega \in \mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$, set $k := k(\omega) := k(p)$ for some shortest U-description p of ω (chosen arbitrarily but fixed for each ω), and suppose $k \le \alpha - K(\alpha) - c_{16}$ for a large c_{16}. We will use Q_k as our model for ω; its complexity can be bounded in two ways:

- Since a self-delimiting description of Q_k (for some prefix-free decoder, not necessarily U) is provided by a shortest U-description of k and the first $k+1$ bits of (the binary expansion of) $\mathbf{\Omega}$, we have
$$K(Q_k) \le k + K(k) + c_{17}. \tag{34}$$
- Another self-delimiting description of Q_k consists of a shortest U-description of α (which gives us both α and $K(\alpha)$), a shortest U-description of $\alpha - K(\alpha) - k$, and the first $k+1$ bits of $\mathbf{\Omega}$; this gives
$$K(Q_k) \le K(\alpha) + K(\alpha - K(\alpha) - k) + k + c_{18} \le \alpha \tag{35}$$
for a large enough c_{16}; the second inequality in (35) follows from the inequality $K(s) \le s/2 + c_{19}$ applied to $s := \alpha - K(\alpha) - k$.

The second item characterizes the complexity of the model Q_k, $K(Q_k) \le \alpha$; let us check that the first item characterizes its quality. Since $Q_k(\{\omega\}) \ge 2^{k-K(\omega)}$, the randomness e-deficiency of ω w.r.t. Q_k is
$$\begin{aligned} -\log Q_k(\{\omega\}) - K(\omega \mid Q_k) &\le \big(K(\omega) - k\big) - \big(K(\omega) - K(Q_k)\big) + c_{20} \\ &= K(Q_k) - k + c_{20} \le K(k) + c_{21}, \end{aligned} \tag{36}$$
where the last inequality follows from (34). Since $\omega \in \mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$, we have $\beta < K(k) + c_{21} \le 2\log(k+1) + c_{22}$, which implies $k > 2^{\beta/2 - c_{23}} - 1$.

We can see that $\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$ is covered by
$$\{\omega \mid k(\omega) > \alpha - K(\alpha) - c_{16}\} \quad \text{and} \quad \{\omega \mid k(\omega) > 2^{\beta/2 - c_{23}} - 1\}. \tag{37}$$
It follows immediately from the definition that
$$\mathbf{m}(\{\omega \mid k(\omega) \ge j\}) \le c_{24} 2^{-j} \quad \text{for all } j \in \mathbb{N}_0, \tag{38}$$
and so (37) gives
$$\mathbf{m}(\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}) \le c_{25} \left(2^{-\alpha + K(\alpha) + c_{16}} + 2^{-2^{\beta/2 - c_{23}}}\right). \qquad \square$$

Let us check that the upper bound of Theorem 7 can be slightly improved to
$$\mathbf{m}(\mathcal{N}^{\mathrm{q}}_{\alpha,\beta}) \le c \left(2^{-\alpha + K(\alpha)} + 2^{-2^{\beta - K(\beta) - c}}\right),$$
assuming $\beta \in \mathbb{N}_0$. The only required change in the proof is to deduce $k > 2^{\beta - K(\beta) - c_{26}} - 1$ from $\beta < K(k) + c_{21}$. To establish this lower bound on k, we argue indirectly. Every $k' \le 2^{\beta - K(\beta) - c_{26}} - 1$ has the following self-delimiting description: a shortest U-description of β (also giving $K(\beta)$), a shortest self-delimiting description of c_{26}, and $\beta - K(\beta) - c_{26}$ bits identifying k' (assuming $c_{26} \in \mathbb{N}$). For a large enough c_{26}, this will contradict $\beta < K(k') + c_{21}$.

Another proof technique that we need is a version of Shen's "first uncovered object". The next lemma presents this technique both for prefix complexity K (which we need now) and for plain complexity C (which we will need later). Remember that the busy beaver function BB is defined by (4) in Sect. 2. The C-counterpart of this function is $\mathrm{BB}^C(s) := \max\{n \mid C(n) \le s\}$.

Lemma 20. *For $s \in \mathbb{N}_0$, set $\omega^*_s := \mathrm{BB}(s) + 1$ and $\omega^{*C}_s := \mathrm{BB}^C(s) + 1$. There is a constant c_{27} such that, for all s,*

- $s < K(\omega_s^*) \le s + c_{27}$ *and every finite set* $\Omega \subseteq \mathcal{X}$ *such that* $K(\Omega) \le s - c_{27}$ *satisfies* $\max \Omega < \omega_s^*$ *(in particular,* $\omega_s^* \notin \Omega$*);*
- $s < C(\omega_s^{*C}) \le s + c_{27}$ *and every finite set* $\Omega \subseteq \mathcal{X}$ *such that* $C(\Omega) \le s - c_{27}$ *satisfies* $\max \Omega < \omega_s^{*C}$ *(in particular,* $\omega_s^{*C} \notin \Omega$*).*

Proof. The K- and C-versions are proved in the same way; let us only discuss the K-version. The first inequality, $K(\omega_s^*) > s$, immediately follows from the definition of BB, and the second, $K(\omega_s^*) \le s + c_{27}$, follows from $K(\mathrm{BB}(s)) \le s$ when $\mathrm{BB}(s) > 0$ (for $\mathrm{BB}(s) = 0$, it suffices to make $c_{27} \ge K(1)$). If a nonempty finite $\Omega \subseteq \mathcal{X}$ satisfies $K(\Omega) \le s - c_{27}$ for a large enough c_{27}, then $K(\max \Omega) \le s$, which implies $\max \Omega \le \mathrm{BB}(s) < \omega_s^*$. □

In particular, all objects $\omega \ge \omega^*_{\alpha + c_{27}}$ belong to $\mathcal{N}_{\alpha,\infty}$, and all objects $\omega \ge \omega^{*C}_{\alpha + c_{27}}$ belong to $\mathcal{N}^C_{\alpha,\infty}$ (in the notation of Sect. 2). Shen's "first uncovered object" [15, Theorem 2(b)] is the smallest object in $\mathcal{N}_{\alpha,\infty}$ or $\mathcal{N}^C_{\alpha,\infty}$; its modifications ω^* and ω^{*C} are more convenient here.

Proof of the lower bound of Theorem 7. If $\omega \notin \mathcal{N}^{\mathrm{q}}_{\alpha,\beta}$, then ω belongs to the support Ω of an elementary probability measure Q with $K(Q) \le \alpha$; we have $K(\Omega) \le \alpha + c_{28}$. Therefore, $\mathcal{N}^{\mathrm{q}}_{\alpha,\infty}$ contains all objects that do not belong to any finite set of prefix complexity at most $\alpha + c_{28}$. By Lemma 20, one of those objects has prefix complexity at most $\alpha + c_{29}$, which gives the lower bound in the theorem. □

Proof of Theorem 8. The lower bound in (10) is proved in the same way, verbatim, as the lower bound in Theorem 7.

The upper bound in (10) has already been proved, essentially, in the proof of the upper bound in Theorem 7. With the unconditional definition, we have, in place of (36),

$$-\log Q_k(\{\omega\}) - K(\omega) \le \big(K(\omega) - k\big) - K(\omega) = -k,$$

and so the covering family (37) only requires the set $\{\omega \mid k(\omega) > \alpha - K(\alpha) - c_{16}\}$.

It remains to prove the second part of the theorem. We will use Gács's result [7] essentially in the form given by Andreev [1, Theorem 2(ii)]: if (x_n, y_n), $n \in \mathbb{N}$, are distinct pairs of nonnegative integers such that $x_n \le y_n$ for all n, the sequence (x_n) is lower semicomputable, the sequence (y_n) is upper semicomputable, and $\sum_n 2^{x_n - y_n} = \infty$, then $\tau(\mathrm{BB}(y_n)) \ge 2^{-x_n}$ for some n. This result trivially extends to n ranging over the integers $n \ge n_0$ for some $n_0 \in \mathbb{N}$. For f as in the statement of the theorem and assuming $n_0 \ge K(1)$, apply Gács's result to

$$x_n := (n - f(n) - 1)^+ \le n + c_{30} =: y_n, \quad n \ge n_0,$$

where c_{30} is a constant that is large enough (in particular, as compared with c_{27} in Lemma 20). We obtain $n \ge n_0$ such that $\tau(\mathrm{BB}(n + c_{30})) \ge 2^{-x_n}$. Since $\mathrm{BB}(n + c_{30}) \ge 1$ and $\tau(t) \le 1 - \mathbf{m}(1) < 1$ for all $t \in \mathbb{N}$, we have $x_n \ne 0$, which implies $x_n = n - f(n) - 1$. By Lemma 20,

$$\mathcal{N}^{\mathrm{qu}}_{\alpha,\beta} \supseteq \{\omega \mid \omega > \mathrm{BB}(\alpha + c_{30})\}$$

holds for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty]$; therefore,

$$\mathbf{m}(\mathcal{N}^{\mathrm{qu}}_{n,\beta}) \ge \tau(\mathrm{BB}(n + c_{30})) \ge 2^{-n + f(n) + 1} > 2^{-n + f(n)}$$

for all $\beta \in [0, \infty]$. Since n_0 can be arbitrarily large, this completes the proof. □

A.5 Proof of Theorem 9

The proof of the lower bound of Theorem 7 relies on the support of elementary probability measures being finite. Computable probability measures can have the whole of $\mathcal{X}$ as their support, of course, but the following lemma shows that their tails can be both light and simple.

Lemma 21. *There is a constant c_{31} such that the following holds for all $\alpha, n \in \mathbb{N}_0$. If a probability measure P on $\mathcal{X}$ has a number π with $K(\pi) \leq \alpha$, there is $N \in \mathbb{N}$ such that $P(\mathcal{X} \setminus [N]) < 2^{-\alpha-n}$ and $K(N) \leq \alpha + K(n) + c_{31}$.*

The parameter n in Lemma 21 describes the trade-off between the lightness and simplicity of tails. Since the upper bound on $K(N)$ involves the complexity of n, a judicious choice of n (using the function μ) will allow a very large value of n and so a very light tail. For a zero tail, as in the case of elementary probability measures, we have no dependence of the lower bound on β, while now we will have light dependence on β.

Proof of Lemma 21. Set $d := \alpha - K(\pi) \geq 0$, and let π^* be a shortest self-delimiting description of π. Given π^* and $n + d$, we can compute $N \in \mathbb{N}$ such that $P(\mathcal{X} \setminus [N]) < 2^{-|\pi^*|-n-d} = 2^{-\alpha-n}$. Therefore,

$$K(N) \leq K(\pi) + K(n+d) + c_{32} \leq K(\pi) + K(n) + d + c_{31} = \alpha + K(n) + c_{31},$$

the second inequality holding for a large enough c_{31}. □

Now we can prove the theorem. Let c_{33} be a large constant (see below), let $n > 2\beta + 2c_{33}$ satisfy $K(n) = \mu(2\beta + 2c_{33})$, and let $\omega^* := \omega^*_{\alpha'}$ be the object of Lemma 20 for $\alpha' := \alpha + K(n) + c_{33}$; therefore, $K(\omega^*) \leq \alpha' + c_{27}$ and, for all $N \geq \omega^*$, $K(N) > \alpha'$.

The object ω^* alone will witness the inequality in Theorem 9. To show that $\omega^* \in \mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}$, we argue indirectly. Suppose ω^* is (α, β)-Quasi-stochastic as witnessed by P and its number π. We then have

$$\begin{aligned} -\log P(\{\omega^*\}) &\leq K(\omega^* \mid \pi) + \beta \leq K(\omega^*) + \beta + c_{34} \leq \alpha' + \beta + c_{35} \\ &= \alpha + K(n) + c_{33} + \beta + c_{35} < \alpha + n; \end{aligned} \tag{39}$$

the last inequality holds since $\beta + c_{33} < n/2$ and, for a large c_{33} (using $n > 2c_{33}$), $K(n) + c_{35} \leq n/2$. Now let N be given by Lemma 21 for P, π, and n. Since $P(\mathcal{X} \setminus [N]) < 2^{-\alpha-n} < P(\{\omega^*\})$, we have $N \geq \omega^*$. On the other hand, $K(N) \leq \alpha + K(n) + c_{31} \leq \alpha'$ for a large c_{33}, which contradicts $N \geq \omega^*$. Therefore, $\omega^* \in \mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}$.

Now we obtain

$$\mathbf{m}(\mathcal{N}^{\mathrm{Q}}_{\alpha,\beta}) \geq \mathbf{m}(\omega^*) \geq c_{36} 2^{-K(\omega^*)} \geq c_{36} 2^{-\alpha'-c_{27}} = c_{36} 2^{-\alpha-\mu(2\beta+2c_{33})-c_{33}-c_{27}}.$$

To complete the proof of the theorem, it remains to notice that $\mu(2\beta + 2c_{33}) \leq \mu(\beta) + c_{37}$.

Apart from c_{33}, every constant in this proof depends only on previously defined constants.

The remark about $\mathcal{N}^{\mathrm{Qu}}_{\alpha,\beta}$ made after Theorem 9 is obvious from (39): we can just remove the expression between the first and second "$\leq$".

A.6 Proofs of Theorems 10, 11, and 5

This section gives (among other things) independent derivations of (15) and (16) in Theorem 11 and of Theorem 5. Alternatively, we could have derived these statements as corollaries of Theorems 13 and 14.

Proof of Theorem 10. Let $\omega \notin \mathcal{N}_{\alpha,\beta}(d)$, as witnessed by an elementary probability measure Q satisfying $K(Q) \le \alpha$ and $d(\omega \mid Q) \le \beta$. We have

$$\begin{aligned} K(\omega) &\le K(Q) - \log Q(\{\omega\}) + c_{38} \le \alpha + F(\omega) + d_F(\omega \mid Q) + c_{38} \\ &\le \alpha + F(\omega) + cd(\omega \mid Q) + c_{39} \le \alpha + F(\omega) + c\beta + c_{39}, \end{aligned} \tag{40}$$

where the third "$\le$" follows from (13). Therefore, $\mathcal{N}_{\alpha,\beta}(d)$ contains all ω satisfying $K(\omega) - F(\omega) > t$, where $t := \lceil \alpha + c\beta + c_{39} \rceil$, and Lemma 18 gives $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d)) \ge c_{40}\tau(t)$. Applying Lemma 3 completes the proof. □

The proof of Theorem 10 shows that Theorem 10 and its C-version (with $\mathcal{N}^C_{\alpha,\beta}(d)$ in place of $\mathcal{N}_{\alpha,\beta}(d)$) can be strengthened by replacing (13) with

$$d_F(\omega \mid Q) \le cd(\omega \mid Q) + cK(Q) + c. \tag{41}$$

Indeed, the chain (40) already contains $K(Q)$, and we can use $K(Q) \le \alpha$ again; for the C-version, this should be replaced by, e.g., $K(Q) \le 2C(Q) + O(1) \le 2\alpha + O(1)$.

Let us state explicitly inequalities that have already been used in narrower contexts (and will be used again in the following proofs). Let δ_ω be the point mass at $\omega \in \mathcal{X}$. The following proposition gives two extreme possibilities for the universal measure of non-quasi-stochastic objects. (All deficiencies considered in this paper satisfy both conditions given in the proposition.)

Proposition 22. *Let d be a deficiency. If $d(\omega \mid Q) = \infty$ whenever $Q(\{\omega\}) = 0$, there exists $c_{41} > 0$ such that $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d)) \ge 2^{-\alpha - c_{41}}$ for all α and β. If $d(\omega \mid \delta_\omega) \le c_{42}$ for all ω, where c_{42} is a constant, there exists $c_{43} > 0$ such that $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d)) \le c_{43}\tau(\alpha)$ for all α and all $\beta \ge c_{42}$.*

Proof. The lower bound is proved in the same way as the lower bound in Theorem 7, and the upper bound is proved in the same way as the upper bound in (2). □

Proof of Theorem 11. We start from the case $\sum_\omega 2^{-F(\omega)} < \infty$, where we have $F \ge K - c_{44}$ and thus $d_F \le d_K + c_{44}$. Therefore,

$$\mathcal{N}_{\alpha,\beta}(d_F) \subseteq \mathcal{N}_{\alpha,\beta - c_{44}}(d_K) = \mathcal{N}^{\mathrm{qu}}_{\alpha,\beta - c_{44}}$$

for all $\beta \ge c_{44}$, and the upper bound in (14) (with the extra condition $\beta \ge c_{44}$) follows from Theorem 8. The lower bound in (14) follows from Proposition 22.

Let us now get rid of the extra condition $\beta \ge c_{44}$ in the upper bound; it suffices to prove that $\mathcal{N}_{\alpha + c_{45},0}(d_F) \subseteq \mathcal{N}_{\alpha,c_{44}}(d_F)$ for some constant c_{45}. To do that, take $\omega \notin \mathcal{N}_{\alpha,c_{44}}(d_F)$, witnessed by a model Q with $K(Q) \le \alpha$ and $d_F(\omega \mid Q) \le c_{44}$. Cut the support of Q, in increasing order, into blocks: each atom of mass greater than $2^{-c_{44}}$ forms a block on its own; the remaining atoms are grouped into maximal consecutive runs of total mass at most $2^{-c_{44}}$. This gives fewer than $2^{c_{44}+2}$ blocks. Then $Q' := Q(\cdot \mid B)$, where B is the block containing ω, will witness $\omega \notin \mathcal{N}_{\alpha + c_{45},0}(d_F)$ for a sufficiently large c_{45}.

Now suppose $\sum_\omega 2^{-F(\omega)} = \infty$. The lower bound in (15) is a special case of Theorem 10, and the upper bound follows from Proposition 22.

Let us prove (16). Combining the lower bounds in (15) and Proposition 22 with Lemma 19, we obtain

$$\begin{aligned}-\log \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) &\le \min\bigl(\alpha, -\log\tau(\max(\alpha,\beta))\bigr) + c_{46}\\ &\le \min\bigl(\alpha, \mu(\max(\alpha,\beta))\bigr) + c_{47} \le \min(\alpha,\mu(\beta)) + 2\log(\alpha+2) + c_{48},\end{aligned}$$

where the last inequality (for a large enough c_{48}) follows from $\mu(\max(\alpha,\beta)) = \max(\mu(\alpha),\mu(\beta)) \le \mu(\alpha)+\mu(\beta)$ and $\mu(\alpha) \le K(\alpha+1) \le 2\log(\alpha+2)+O(1)$.

For the opposite inequality

$$-\log \mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) \ge \min(\alpha,\mu(\beta)) - c\log(\alpha+2) \tag{42}$$

we use the uniform probability measures $U_{[2^b]}$, $b\in\mathbb{N}_0$, on $[2^b]$; we have $K(U_{[2^b]}) \le K(b)+c_{49}$, and $d_F(\omega \mid U_{[2^b]}) \le b$ for all $\omega\in[2^b]$. Let b^* be the largest integer $b\in[0,\beta]$ satisfying $K(b)\le\alpha-c_{49}$; if there is no such b, then $\alpha < K(0)+c_{49}$, and (42) holds for a large enough c. All $\omega\in[2^{b^*}]$ are (α,β)-quasi-stochastic w.r.t. d_F, and so $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) \le \tau(2^{b^*}) \le \tau(b^*)$. By the definition of b^*, all integers $n\in(b^*,\beta]$ satisfy $K(n)>\alpha-c_{49}$, and all $n>\beta$ satisfy $K(n)\ge\mu(\beta)$; therefore, $\mu(b^*)\ge\min(\alpha-c_{49},\mu(\beta))$. It remains to apply Andreev's estimate $-\log\tau \ge \mu - O(\log\mu)$ (see (6)) to $\tau(b^*)$, taking into account that the function $x\mapsto x-c\log(x+2)$ is increasing for large x. □

Proof of Theorem 5. The proof of (16) carries over verbatim after the following three replacements:

- the lower bound in (15) is replaced by the lower bound in (2);
- the lower bound $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) \ge 2^{-\alpha-c}$ is replaced by $\mathbf{m}(\mathcal{N}_{\alpha,\beta}) \ge 2^{-\alpha-c}$, which follows from Lemma 20: for $\omega^* := \omega^*_{\alpha+c_{27}} \in \mathcal{N}_{\alpha,\infty} \subseteq \mathcal{N}_{\alpha,\beta}$ we have $K(\omega^*) \le \alpha + 2c_{27}$, and so $\omega^*\in\mathcal{N}_{\alpha,\beta}$ and
$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}) \ge \mathbf{m}(\omega^*) \ge c_{50}2^{-\alpha-2c_{27}};$$
- the uniform probability measures $U_{[2^b]}$ on $[2^b]$ are replaced by the finite sets $[2^b]$. □

A.7 Proof of Theorem 13

We prove the upper and lower bounds separately.

Upper bounds

We will use the slices $k(\omega)$ and the elementary probability measures Q_k defined in the proof of the upper bound of Theorem 7 in Sect. A.4. Besides (35), (36), (38), and

$$Q_{k(\omega)}(\{\omega\}) \ge 2^{k(\omega)-K(\omega)} \tag{43}$$

(noticed just before (36)), we need one further property of the slices. Since the first $k+1$ bits of $\mathbf{\Omega}$ form a plain description of Q_k (their number giving k), we have

$$C(Q_k) \le k + c_{51}. \tag{44}$$

For a deficiency d and $\beta\in[0,\infty)$, set

$$B_\beta(d) := \{\omega \mid d(\omega \mid Q_{k(\omega)}) > \beta\};$$

these are the ω for which $Q_{k(\omega)}$ does not serve as a good model. Let $\omega \notin B_\beta(d)$. If $k(\omega) \le \alpha - K(\alpha) - c_{16}$, then $K(Q_{k(\omega)}) \le \alpha$ by (35), and so $\omega \notin \mathcal{N}_{\alpha,\beta}(d)$; if $k(\omega) \le \alpha - c_{51}$, then $C(Q_{k(\omega)}) \le \alpha$ by (44), and so $\omega \notin \mathcal{N}^C_{\alpha,\beta}(d)$. By (38), there is a constant c_{52} such that, for all d, α, and β,

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d)) \le c_{52}\left(2^{-\alpha+K(\alpha)} + \mathbf{m}(B_\beta(d))\right), \tag{45}$$

$$\mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d)) \le c_{52}\left(2^{-\alpha} + \mathbf{m}(B_\beta(d))\right). \tag{46}$$

It remains to bound $\mathbf{m}(B_\beta(d))$ for our four deficiencies.

Since $d_K(\omega \mid Q_{k(\omega)}) \le -k(\omega) \le 0$ (see (43)), the set $B_\beta(d_K)$ is empty, which gives us the upper bounds in (20) and (24).

By (36) (which holds for every ω satisfying $k(\omega) = k$), every $\omega \in B_\beta(d_{\mathrm{e}})$ satisfies $\beta < K(k(\omega)) + c_{21} \le 2\log(k(\omega)+1) + c_{53}$, which implies $k(\omega) + 1 > 2^{(\beta - c_{53})/2}$; therefore, (38) gives $\mathbf{m}(B_\beta(d_{\mathrm{e}})) = O(2^{-2^{(\beta-c_{53})/2}})$. This gives the upper bounds in (18) and (22).

Next let $\omega \in B_\beta(d_{\mathrm{p}})$ and $k := k(\omega)$. Combining (36) and

$$d_{\mathrm{e}}(\omega \mid Q) \ge d_{\mathrm{p}}(\omega \mid Q) - 2\log(d_{\mathrm{p}}(\omega \mid Q) + 2) - c_{54}$$

(see the argument in [12, Proposition 1]), we obtain

$$K(k) \ge d_{\mathrm{e}}(\omega \mid Q_k) - c_{55} \ge \beta - 2\log(\beta+2) - c_{56}.$$

On the other hand, if $k < 2^{\beta/2}$, a shortest U-description of $j := \lceil \beta/2 \rceil$ followed by the j-bit binary representation of k is a self-delimiting description of k, and so

$$K(k) \le j + K(j) + c_{57} \le \beta/2 + 2\log(\beta+2) + c_{58}.$$

The two bounds on $K(k)$ are incompatible when $\beta \ge c_{59}$ (for a large enough c_{59}). Therefore, for $\beta \ge c_{59}$, $B_\beta(d_{\mathrm{p}}) \subseteq \{\omega \mid k(\omega) \ge 2^{\beta/2}\}$, and (38) gives $\mathbf{m}(B_\beta(d_{\mathrm{p}})) = O(2^{-2^{\beta/2}})$. This gives the upper bounds in (19) and (23) (for $\beta < c_{59}$, they are trivial for a large enough c).

Finally, by (43) and $F \ge 0$,

$$d_F(\omega \mid Q_{k(\omega)}) \le K(\omega) - k(\omega) - F(\omega) \le K(\omega),$$

and so $B_\beta(d_F) \subseteq \{\omega \mid K(\omega) > \beta\}$, whose measure is $\sigma(\beta) = O(\tau(\beta))$ by Lemma 4. By (45) and (46),

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) = O(2^{-\alpha+K(\alpha)} + \tau(\beta)) \quad \text{and} \quad \mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d_F)) = O(2^{-\alpha} + \tau(\beta)).$$

On the other hand, since $d_F(\omega \mid \delta_\omega) = -F(\omega) \le 0$, Proposition 22 gives $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d_F)) = O(\tau(\alpha))$. Since $C \le K + c_{60}$, we have $\mathcal{N}^C_{\alpha,\beta}(d) \subseteq \mathcal{N}_{\alpha-c_{60},\beta}(d)$ for $\alpha \ge c_{60}$, and so Lemma 3 gives $\mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d_F)) = O(\tau(\alpha))$ as well. Combining these bounds and using

$$\min(\tau(\alpha), x + \tau(\beta)) \le x + \tau(\max(\alpha,\beta)) \tag{47}$$

for $x \ge 0$, we obtain the upper bounds in (21) and (25).

Lower bounds

All four deficiencies satisfy $d(\omega \mid Q) = \infty$ whenever $Q(\{\omega\}) = 0$. Therefore, Proposition 22 gives the lower bound $2^{-\alpha}/c$ in (18)–(21). Similarly, every $\omega \notin \mathcal{N}^C_{\alpha,\beta}(d)$ belongs to the support of an elementary probability measure Q with $C(Q) \le \alpha$, which is a finite set of plain complexity at most

$\alpha' := \alpha + c_{61}$. By Lemma 20, the object $\omega^* := \omega^{*C}_{\alpha'+c_{27}} \in \mathcal{N}^C_{\alpha',\infty}$ has $C(\omega^*)$ within an additive constant of α', and so it satisfies

$$K(\omega^*) \le C(\omega^*) + K(C(\omega^*)) + O(1) \le \alpha' + K(\alpha') + O(1) \le \alpha + K(\alpha) + O(1) \tag{48}$$

(for the first inequality, notice that a shortest self-delimiting description of $C(\omega^*)$ followed by a shortest plain description of ω^* is a self-delimiting description of ω^*). This gives the lower bound $2^{-\alpha-K(\alpha)}/c$ in (22)–(25). Finally, the terms $\tau(\max(\alpha,\beta))$ in the lower bounds in (21) and (25) are given by Theorem 10 (with $d := d_F$, for which (13) is trivial) and its C-version stated after it.

A.8 Proof of Theorem 14

The proof of the upper bound in Theorem 14 will use the following lemma. Let us say that $I \subseteq \mathbb{N}$ is an *interval* if $I = \{n \mid a < n \le b\}$ for some $a, b \in \mathbb{N}_0$.

Lemma 23. *There is a constant c_{62} such that, for all $s \in \mathbb{N}_0$ and all intervals $I \subseteq \mathbb{N}$ satisfying $K(n) > s$ for all $n \in I$,*

$$\mathbf{m}(I) \le 2^{-s+c_{62}\log(s+2)}. \tag{49}$$

Proof. Fix a non-empty interval $I = \{n \mid a < n \le b\}$ (for $I = \emptyset$, (49) is trivial). We use the slices and the elementary probability measures Q_k defined in the proof of the upper bound of Theorem 7. It is clear that

$$\mathbf{m}(I) \le c_{63} \sum_{k=0}^{\infty} 2^{-k} Q_k(I). \tag{50}$$

Let us estimate $Q_k(I)$ for $k \in \mathbb{N}_0$. Suppose $Q_k(I) > 2^{-r}$ for $r \in \mathbb{N}$; we will find a simple $y \in I$. Set $G(x) := Q_k([x])$ for $x \in \mathbb{N}_0$ (with $G(0) := 0$) and

$$y := \min\{x \mid G(x) > \lceil 2^{r+1} G(a) \rceil 2^{-r-1}\}$$

(so that y is the first point at which G exceeds $G(a)$ rounded upwards to $r+1$ binary places). It is clear that $y \in I$, and we can compute y from Q_k, r, and $r+2$ bits; therefore, (34) gives

$$s < K(y) \le k + K(k) + r + K(r) + c_{64}.$$

Therefore, $Q_k(I) \le 2^{-r}$ for all $r \in \mathbb{N}$ satisfying $r + K(r) \le s - k - K(k) - c_{64}$, and taking the largest such r gives

$$Q_k(I) \le 2^{-s+k+K(k)+2\log(s+2)+c_{65}} \tag{51}$$

for all k (for a large enough c_{65}). The case where such r does not exist should be considered separately; (51) is then trivial (again for a large c_{65}). Plugging (51) for $k \le s$ and $Q_k(I) \le 1$ for $k > s$ into (50) and using $K(k) \le 2\log(k+2) + O(1)$, we obtain

$$\mathbf{m}(I) \le c_{63} \left(\sum_{k=0}^{s} 2^{-s+K(k)+2\log(s+2)+c_{65}} + \sum_{k=s+1}^{\infty} 2^{-k} \right) = O\left(2^{-s}(s+2)^5\right),$$

and it remains to take a large enough c_{62}. □

Lemma 23 may be said to cover a middle ground between (6) (covering intervals of the form (a,∞)) and Epstein and Levin's results (covering arbitrary $I \subseteq \mathbb{N}$).

Proof of Theorem 14. Let $d \in \{d_{\mathrm{e}}, d_{\mathrm{p}}, d_K, d_F\}$. We start from the lower bounds. Since $d^{\mathrm{set}} \geq d$, we have $\mathcal{N}_{\alpha,\beta}(d) \subseteq \mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})$ and $\mathcal{N}^C_{\alpha,\beta}(d) \subseteq \mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}})$, and so Theorem 13 gives the terms $2^{-\alpha}$ in (26) and $2^{-\alpha-K(\alpha)}$ in (27), as well as the terms $\tau(\max(\alpha,\beta))$ for $d = d_F$. For $d \in \{d_{\mathrm{e}}, d_{\mathrm{p}}\}$, the terms $\tau(\max(\alpha,\beta))$ are given by Theorem 10 with $F := C$ and its C-version (stated right after Theorem 10): (13) holds in the form $d_C \leq 2d^{\mathrm{set}}_{\mathrm{e}} + O(1)$ by (32) and in the form $d_C \leq d^{\mathrm{set}}_{\mathrm{p}} + O(1)$ by (17). For $d = d_K$, a similar argument is applicable if we replace (13) in Theorem 10 by (41): indeed, $d_C \leq cd^{\mathrm{set}}_K + cK(Q) + c$ follows from

$$d_C(\omega \mid U_\Omega) \leq 2d^{\mathrm{set}}_{\mathrm{e}}(\omega \mid U_\Omega) + O(1) \leq 2d^{\mathrm{set}}_K(\omega \mid U_\Omega) + 2K(\Omega) + O(1).$$

Now we prove the upper bound (28). Since the point mass δ_ω is the uniform probability measure on $\{\omega\}$, we have $d^{\mathrm{set}}(\omega \mid \delta_\omega) = d(\omega \mid \delta_\omega) \leq 0$ for all ω, and so Proposition 22 gives $\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})) = O(\tau(\alpha))$; as in the proof of Theorem 13, this implies $\mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}})) = O(\tau(\alpha))$. On the other hand, for all $b \in \mathbb{N}_0$, $K(U_{[2^b]}) \leq K(b) + c_{66}$ and $d^{\mathrm{set}}(\omega \mid U_{[2^b]}) \leq b + c_{67}$ for all $\omega \in [2^b]$ (e.g., $d^{\mathrm{set}}_{\mathrm{p}}(\omega \mid U_{[2^b]}) = b - C(\omega \mid [2^b]) + O(1)$). Let b^* be the largest $b \in \mathbb{N}_0$ satisfying $b \leq \beta - c_{67}$ and $K(b) \leq \alpha - c_{66}$ (if there is no such b, then $\beta < c_{67}$ or $\alpha < K(0) + c_{66}$, and (28) follows from the bound $O(\tau(\alpha))$ or is trivial for a large enough c). All $\omega \in [2^{b^*}]$ are (α,β)-quasi-stochastic w.r.t. d^{set} (with $U_{[2^{b^*}]}$ as the model), and so

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})) \leq \tau(2^{b^*}) \leq \tau(b^*) = \mathbf{m}(I) + \tau(\beta - c_{67}),$$

where $I := \{n \mid b^* < n \leq \beta - c_{67}\}$. By the definition of b^*, all $n \in I$ satisfy $K(n) > \alpha - c_{66}$, and so Lemma 23 gives $\mathbf{m}(I) \leq 2^{-\alpha + c_{68}\log(\alpha+2)}$ for a large enough c_{68}; besides, $\tau(\beta - c_{67}) = O(\tau(\beta))$ by Lemma 3. Since $\mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}}) \subseteq \mathcal{N}_{\alpha-c_{60},\beta}(d^{\mathrm{set}})$ (see the proof of Theorem 13), the resulting bound

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})) \leq 2^{-\alpha+c_{68}\log(\alpha+2)} + O(\tau(\beta))$$

also holds for $\mathbf{m}(\mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}}))$ (with a larger c_{68}). Combining these bounds with the earlier bound $O(\tau(\alpha))$ and again using (47), we obtain (28). □

A.9 Proof of Theorem 15

Let B be one of the sets $\{\omega \mid K(\omega) > k\}$ and $\{\omega \mid C(\omega) > k\}$, and let $\mathcal{N}$ be one of the sets covered by Theorems 13 and 14. We start from the upper bounds. For $B = \{\omega \mid K(\omega) > k\}$, we have $\mathbf{m}(B) = \sigma(k) = O(\tau(k))$ by (3). In combination with Lemma 3, this also gives $\mathbf{m}(B) = O(\tau(k))$ for $B = \{\omega \mid C(\omega) > k\}$. Therefore, each upper bound $\mathbf{m}(\mathcal{N}) \leq cU$ implies

$$\mathbf{m}(\mathcal{N} \cap B) \leq \min\bigl(\mathbf{m}(\mathcal{N}), \mathbf{m}(B)\bigr) = O\bigl(\min(U, \tau(k))\bigr).$$

As for the lower bounds, we first unify our discussion of different ways of measuring the complexity of models by defining the *floor* to be $2^{-\alpha}$ for the sets $\mathcal{N}_{\alpha,\beta}(\cdot)$ and to be $2^{-\alpha-K(\alpha)}$ for the sets $\mathcal{N}^C_{\alpha,\beta}(\cdot)$. In each lower bound $\mathbf{m}(\mathcal{N}) \geq L/c$ in Theorems 13 and 14, L consists of one or two addends: the floor and, in 10 of 16 cases, $\tau(\max(\alpha,\beta))$. Since $\min(x+y,z) \leq \min(x,z) + \min(y,z)$ whenever $x, y, z \geq 0$, it suffices to prove $\mathbf{m}(\mathcal{N} \cap B) \geq \min(x, \tau(k))/c$ for each addend of L separately.

First we deal with the floors. Let $\mathcal{N}$ be $\mathcal{N}_{\alpha,\beta}(d)$ or $\mathcal{N}_{\alpha,\beta}(d^{\mathrm{set}})$, where $d \in \{d_{\mathrm{e}}, d_{\mathrm{p}}, d_K, d_F\}$ (the proof for the $\mathcal{N}^C$ versions will be given later, with some preparatory work for $\mathcal{N}^C$ already done in the proof for $\mathcal{N}$). The support Ω of an elementary probability measure Q satisfies $K(\Omega) \leq K(Q) + c_{69}$, and so, by Lemma 20 with $s := \alpha + c_{70}$ for a large c_{70}, no object in the tail

$$T := \{\omega \mid \omega > \mathrm{BB}(s)\}$$

belongs to the support of an elementary probability measure Q with $K(Q) \le \alpha$. Since $d(\omega \mid Q) = d^{\mathrm{set}}(\omega \mid Q) = \infty$ whenever $Q(\{\omega\}) = 0$, we have $T \subseteq \mathcal{N}$.

For $n \in \mathbb{N}$, let κ_n be the uniform probability measure on the 2^n numbers $n2^n, \dots, (n+1)2^n - 1$. Since at most 2^{n-1} objects ω satisfy $K(\omega) \le n-1$ and fewer than 2^{n-1} objects satisfy $C(\omega) \le n-2$, κ_n gives probability at least $1/2$ both to $\{\omega \mid K(\omega) \ge n\}$ and to $\{\omega \mid C(\omega) \ge n-1\}$. For $n \in T$, κ_n is concentrated on T, and so $\kappa_n(T \cap B) \ge 1/2$ whenever $n > \max(\mathrm{BB}(s), k+1)$, and Lemma 17 gives

$$\mathbf{m}(\mathcal{N} \cap B) \ge \mathbf{m}(T \cap B) \ge \frac{c_\kappa}{2} \tau\bigl(\max(\mathrm{BB}(s), k+1)\bigr) = \frac{c_\kappa}{2} \min\bigl(\tau(\mathrm{BB}(s)), \tau(k+1)\bigr). \tag{52}$$

It remains to notice that $\tau(\mathrm{BB}(s)) \ge \mathbf{m}(\omega^*_s) \ge 2^{-s-c_{71}} \ge 2^{-\alpha-c_{72}}$ by Lemma 20 and that $\tau(k+1) \ge c_{73}\tau(k)$ by Lemma 3.

Now let us deal with $\mathcal{N}^C_{\alpha,\beta}(d)$ and $\mathcal{N}^C_{\alpha,\beta}(d^{\mathrm{set}})$. The argument is the same, except that now the support Ω of Q satisfies $C(\Omega) \le C(Q) + c_{74}$, and so, by Lemma 20, T can be replaced by $\{\omega \mid \omega > \mathrm{BB}^C(s')\}$, where $s' := \alpha' + c_{27}$ and $\alpha' := \alpha + c_{74}$. In (52), $\mathrm{BB}(s)$ is then replaced by $\mathrm{BB}^C(s')$, and it remains to notice that

$$\tau(\mathrm{BB}^C(s')) \ge \mathbf{m}(\omega^{*C}_{s'}) \ge c_{75} 2^{-\alpha - K(\alpha)},$$

since

$$K(\omega^{*C}_{s'}) \le \alpha' + K(\alpha') + O(1) \le \alpha + K(\alpha) + O(1)$$

(see (48) in the proof of the lower bounds in Theorem 13).

It remains to deal with the addend $\tau(\max(\alpha, \beta))$, which appears in the lower bounds for d_F in Theorem 13 and for all four d^{set} in Theorem 14. There it was obtained from Theorem 10 or its C-version, with (13) or (41), applied to some F (perhaps $F := C$). The chain (40), combined with $K(Q) \le 2C(Q) + O(1)$ for the C-versions, shows that, in all these cases, $\mathcal{N} \supseteq \{\omega \mid K(\omega) - F(\omega) > t\}$, where $t := \lceil c_{76}(\alpha + \beta + 1) \rceil$ for a large enough c_{76}. Set $k' := k$ if $B = \{\omega \mid K(\omega) > k\}$ and $k' := k + \lceil 2\log(k+2) \rceil + c_{77}$ if $B = \{\omega \mid C(\omega) > k\}$; in the latter case, $K(\omega) > k'$ implies $C(\omega) > k$, provided c_{77} is large enough. Since $F \ge 0$,

$$\{\omega \mid K(\omega) - F(\omega) > \max(t, k')\} \subseteq \mathcal{N} \cap \{\omega \mid K(\omega) > k'\} \subseteq \mathcal{N} \cap B,$$

and Lemma 18 gives

$$\mathbf{m}(\mathcal{N} \cap B) \ge c_{78} \tau\bigl(\max(t, k')\bigr) = c_{78} \min\bigl(\tau(t), \tau(k')\bigr).$$

Finally, Lemma 3 gives $\tau(k') \ge c_{79}\tau(k)$ and $\tau(t) \ge c_{80}\tau(\max(\alpha, \beta))$ (remember that, according to the definition, $\tau(x) = \tau(\lfloor x \rfloor)$ for $x \ge 0$).

B Kolmogorov's questions

This appendix gives the historical motivation for this paper. In the next paragraph I reproduce Kolmogorov's questions posed in his talk on 26 November 1981 and briefly recorded by Shen [14, note 5.12].

A finite object ω is (α, β)-stochastic if there exists a finite set Ω such that

- $\omega \in \Omega$;
- $K(\Omega) \le \alpha$;

- $K(\omega) \geq \log|\Omega| - \beta$.

Let $N(k, \alpha, \beta)$ be the number of (α, β)-stochastic objects of complexity at most k. Typical cases are $\alpha, \beta = \epsilon k$ and $\alpha, \beta = \sqrt{k}$. What can we say about the asymptotics of $N(k, \alpha, \beta)/2^k$ as $k \to \infty$? Is it true that $N(k, \alpha, \beta)/2^k \to 0$, if k increases and $\alpha/k, \beta/k \to 0$?

It is very likely that by K Kolmogorov meant plain complexity, nowadays usually denoted by C; it appears that he never used prefix complexity. It is also likely that there is a typo in the definition of $N(k, \alpha, \beta)$ and in fact it means the number of (α, β)-nonstochastic objects of complexity at most k. Therefore, in this appendix we set

$$N(k, \alpha, \beta) := |D_k \setminus \mathcal{S}_{\alpha,\beta}| = \left|D_k \cap \mathcal{N}^{C\mathrm{u}}_{\alpha,\beta}\right|,$$

where $\mathcal{N}^{C\mathrm{u}}_{\alpha,\beta}$ was introduced in Sect. 4 and

$$\begin{aligned} D_k &:= \{\omega \in \mathcal{X} \mid C(\omega) \leq k\}, \\ \mathcal{S}_{\alpha,\beta} &:= \left\{\omega \in \mathcal{X} \mid \exists \Omega \subseteq \mathcal{X} : |\Omega| < \infty, \omega \in \Omega, C(\Omega) \leq \alpha, C(\omega) \geq \log|\Omega| - \beta\right\}. \end{aligned} \tag{53}$$

Here we refer only to the elements of $\mathcal{S}_{\alpha,\beta}$ as (α, β)-stochastic (unlike in the rest of the paper, where stochasticity was also understood in other related senses).

Remark 24. Reading "stochastic" literally, i.e., replacing $N(k, \alpha, \beta)$ with

$$|D_k \cap \mathcal{S}_{\alpha,\beta}| = |D_k| - N(k, \alpha, \beta),$$

makes both of Kolmogorov's questions degenerate. Let c be a constant such that $C(\omega) \leq n + c$ for all n and all $\omega \in [2^n]$. For $n := k - c$, the set $\Omega := [2^n]$ is a subset of D_k with $C(\Omega) \leq \log k + O(1)$, and more than half of its elements satisfy $C(\omega) \geq n - 1 = \log|\Omega| - 1$. Hence $|D_k \cap \mathcal{S}_{\alpha,\beta}| \geq 2^{k-c-1}$ as soon as $\beta \geq 1$ and $\alpha \geq \log k + O(1)$, and the answer to the last question would be negative already in the typical case $\alpha = \beta = \sqrt{k}$. Moreover, by (54) below, in both typical cases $|D_k \cap \mathcal{S}_{\alpha,\beta}|/2^k$ differs by $o(1)$ from $|D_k|/2^k \in [2^{-c}, 2)$, which involves neither α nor β; so the first question would not be about stochasticity at all. With "nonstochastic", both questions become substantive: see Theorem 25 and Corollary 26.

Two other facts point the same way. Shen's 1983 paper, presented to *Doklady* by Kolmogorov himself, bounds the number of nonstochastic objects [15, Theorem 3], and V'yugin, another participant in the seminar, recalls Kolmogorov asking in the talk whether nonstochastic strings exist and, if so, how many of them there are among the strings of a given length [21, Sect. 2]. (Another discrepancy between Shen's record and V'yugin's recollection is that V'yugin mentions lengths of strings while Shen's record is structure-free, but contemporaneous records are usually more reliable.) Slips in the record are not surprising: its definition of stochasticity contains another one, $K(\Omega)$ in place of $\log|\Omega|$, flagged by the authors [14, note 5.12] and corrected above.

Kolmogorov used the condition $C(\omega) \geq \log|\Omega| - \beta$ rather than the more standard $C(\omega \mid \Omega) \geq \log|\Omega| - \beta$ or $K(\omega \mid \Omega) \geq \log|\Omega| - \beta$, which were the principal ones used in the main paper. The suggestion to measure the randomness deficiency of ω as an element of Ω by $\log|\Omega| - C(\omega \mid \Omega)$ rather than $\log|\Omega| - C(\omega)$ is due to Per Martin-Löf (personal communication; see also [11, Theorem on p. 607]), and it leads to an elegant theory: see, e.g., Propositions 11 and 13 in the summary [19, Appendix 1].

B.1 Some answers

The following result answers Kolmogorov's specific questions; its core part is given as (56) after its statement.

Theorem 25. *There is a positive constant c such that, for all $k \in \mathbb{N}$, $\alpha \in \mathbb{N}_0$, and $\beta \in [0, \infty]$,*

$$N(k, \alpha, \beta) \le 2^{k-\alpha+c\log(k+1)} \tag{54}$$

and, provided $\alpha + c \le k$,

$$N(k, \alpha, \beta) \ge 2^{k-\alpha-K(\alpha)-c}. \tag{55}$$

In particular,

$$N(k, \alpha, \beta) = 2^{k-\alpha+O(\log(k+1))} \tag{56}$$

uniformly in $\beta \in [0, \infty]$ whenever $\alpha + c \le k$. The form of (56) is familiar: the fraction of non-stochastic objects among the objects of complexity at most k is about $2^{-\alpha}$.

The theorem will be proved in the next section, and now we will use it to answer Kolmogorov's questions. The following corollary uses $N(k, \alpha, \beta) = N(k, \lfloor\alpha\rfloor, \beta)$ for α that are not necessarily integer.

Corollary 26. *For Kolmogorov's two typical cases $\alpha, \beta = \epsilon k$ (with $\epsilon \in (0, 1)$ fixed) and $\alpha, \beta = \sqrt{k}$ we have, for large k,*

$$\frac{N(k, \epsilon k, \epsilon k)}{2^k} = 2^{-\epsilon k + O(\log k)}, \qquad \frac{N(k, \sqrt{k}, \sqrt{k})}{2^k} = 2^{-\sqrt{k} + O(\log k)},$$

so in both cases the ratio tends to 0 faster than any power of k.

Proof. In both cases $\alpha + c \le k$ for large k, so (56) applies. □

Kolmogorov's last question appears to contain another slip. A negative answer is provided already by a fixed (arbitrarily large) α. This can be seen both from (55) and directly, without using the theorem. Indeed, for a fixed α the family $\{\Omega \mid C(\Omega) \le \alpha\}$ is finite (of size at most $2^{\alpha+1}$); let F_α be its union, a finite set. Every (α, β)-stochastic object is in F_α, whatever β may be, so $N(k, \alpha, \beta) \ge |D_k| - |F_\alpha| \ge 2^{k-O(1)} - O(1)$ and, therefore,

$$\liminf_{k\to\infty} \frac{N(k, \alpha, \beta)}{2^k} \ge c$$

for a positive constant c. Probably Kolmogorov meant the question "when is it true that $N(k, \alpha, \beta)/2^k \to 0$?" According to (54), "$\log k = o(\alpha)$" is a sufficient condition. This condition, however, is far from being necessary (e.g., by (54), it suffices to require $\alpha - c\log k \to \infty$).

B.2 Proof of Theorem 25

We prove the two parts of Theorem 25, (54) and (55), separately, starting with (54). Since $\mathcal{S}_{\alpha,\beta}$ is increasing in (both α and) β, $N(k, \alpha, \beta)$ is decreasing in β. Therefore, it suffices to prove (54) for $\beta = 0$ and (55) for $\beta = \infty$.

Proof of (54). This proof illustrates the technique of "bounded complexity lists" in a simple setting. Fix a jointly computable way of enumerating all elements of D_k (defined by (53)) given k (running the optimal plain decoder on all inputs of length at most k in parallel and listing the outputs in the order in which they first appear). Cutting the enumeration into consecutive blocks of length 2^{k-a} and taking the block containing ω means, intuitively, spending about a bits on the block index and leaving $k - a$ bits of noise inside the block.

The idea behind the technique is that the completed blocks consist, roughly, of stochastic objects. Namely, there is a constant c_{81} such that the following holds for each $k \in \mathbb{N}_0$ (the *level*) and each

$a \in \{0, \dots, k\}$ (the *parameter*). Let $\omega \in D_k$ lie in a complete block of size 2^{k-a} (i.e., a block that the enumeration of D_k lists in full). Then there is a finite $\Omega \ni \omega$ with

$$\log |\Omega| = k - a, \qquad C(\Omega) \le a + K(k) + c_{81}. \tag{57}$$

Indeed, let Ω be the block containing ω (so that the first equality in (57) follows from $|\Omega| = 2^{k-a}$). As the block is complete, Ω is computable from k, a, and the ordinal number $i < 2^{a+1}$ of the block. Its description may consist of a self-delimiting description of k followed by exactly $a+1$ further bits for i (those further bits also determining a).

Choose $k \in \mathbb{N}$ and set $a := \alpha - \lfloor c_{82} \log(k+1) \rfloor$ for a large c_{82} as compared with c_{81}; make c (in Theorem 25) large as compared with c_{82}. If $a < 0$ then $\alpha < \lfloor c_{82} \log(k+1) \rfloor$, and the asserted bound is at least $2^{k+\log(k+1)} \ge 2^{k+1} > |D_k|$, so there is nothing to prove; assume $a \ge 0$.

Let $\omega \in D_k$ and $k' := C(\omega) \le k$. If $k' < a$, apply (57) with the level k' and the parameter k': the block size is 1, so $\Omega = \{\omega\}$,

$$C(\Omega) \le k' + K(k') + c_{81} \le \alpha,$$

and $\log |\Omega| = 0 \le C(\omega)$; therefore, ω is $(\alpha, 0)$-stochastic. If $k' \ge a$ and the block of ω at level k' with parameter a is complete, (57) gives

$$C(\Omega) \le a + K(k') + c_{81} \le \alpha$$

and $\log |\Omega| = k' - a \le k' = C(\omega)$, so again ω is $(\alpha, 0)$-stochastic.

Hence every $(\alpha, 0)$-nonstochastic $\omega \in D_k$ satisfies $k' = C(\omega) \ge a$ and lies in the last, incomplete block at its own level k', and there are fewer than $2^{k'-a}$ such ω. Summing over $k' \in \{a, \dots, k\}$,

$$N(k, \alpha, 0) \le \sum_{k'=a}^{k} 2^{k'-a} < 2^{k-a+1} = 2^{k-\alpha+\lfloor c_{82} \log(k+1) \rfloor + 1} \le 2^{k-\alpha+c \log(k+1)}. \qquad \square$$

Proof of (55). Set $\omega^* := \omega^{*C}_{\alpha+c_{27}}$. By Lemma 20, all objects $\omega \ge \omega^*$ lie outside $\mathcal{S}_{\alpha,\infty}$, and $\alpha + c_{27} < C(\omega^*) \le \alpha + 2c_{27}$. We can see that $C(\omega^*)$ differs from α by at most a constant, which implies

$$K(\omega^*) \le C(\omega^*) + K(C(\omega^*)) + c_{83} \le \alpha + K(\alpha) + c_{84} \tag{58}$$

(as in (48)).

Let $\Delta := k - \alpha - K(\alpha) - c_{85}$, where c_{85} is a large integer (see below). If $\Delta < 0$, it suffices to notice that $\omega^* \in D_k \setminus \mathcal{S}_{\alpha,\infty}$, since $C(\omega^*) \le \alpha + 2c_{27} \le k$ (using the assumption $\alpha + c \le k$ for $c \ge 2c_{27}$); therefore, $N(k, \alpha, \infty) \ge 1 > 2^{\Delta}$. Suppose now that $\Delta \ge 0$, and consider the 2^{Δ} objects $\omega^* + i$, $i \in \{0, \dots, 2^{\Delta} - 1\}$, all of which lie outside $\mathcal{S}_{\alpha,\infty}$. Since a shortest self-delimiting description of ω^* followed by the binary representation of i padded with leading 0s to exactly Δ bits is a plain description of $\omega^* + i$, (58) gives

$$C(\omega^* + i) \le K(\omega^*) + \Delta + c_{86} \le \alpha + K(\alpha) + \Delta + c_{87} = k - c_{85} + c_{87} \le k,$$

provided c_{85} is large enough. Therefore, all objects $\omega^* + i$ belong to $D_k \setminus \mathcal{S}_{\alpha,\infty}$, and so we have $N(k, \alpha, \infty) \ge 2^{\Delta}$. $\square$

The "in particular" bit after the theorem follows from $K(\alpha) \le 2 \log(\alpha+2) + c_{88} \le 2 \log(k+1) + c_{89}$.

Acknowledgments

This paper was conceived during my work on the V'yugin Special Issue [8]. Vladimir V'yugin, one of my teachers in the algorithmic theory of randomness, was a key contributor to the literature on Kolmogorov's (α, β)-stochasticity.

Many thanks to Sasha Shen for several discussions of Kolmogorov's definition of stochastic objects, and also for his record of Kolmogorov's talk 45 years ago.

AI methodology statement

I acknowledge the use of Claude Opus 5–5.5 and Fable 5.1 in exploring proof ideas and developing proofs, which I reviewed carefully, and in checking the paper. I take full responsibility for this paper's claims and statements, including mathematical statements and their proofs. More generally, I have verified the correctness and originality of all content, including references. See Sect. 1 and Appendix A for further details.

References

[1] Mikhail Andreev. Busy beavers and Kolmogorov complexity. In Arnold Beckmann, Laurent Bienvenu, and Nataša Jonoska, editors, *Pursuit of the Universal, Proceedings of the Twelfth Conference on Computability in Europe, CiE 2016*, volume 9709 of *Lecture Notes in Computer Science*, pages 195–204, 2016. Also published as arXiv report (more accessible, and so all references in the paper are to it) arXiv:1703.05170 [cs.CC] (March 2017).

[2] Bruno Bauwens. m-sophistication. Technical Report arXiv:1001.4687 [cs.CC], arXiv.org e-Print archive, January 2010.

[3] Laurent Bienvenu, Glenn Shafer, and Alexander Shen. Andrei Kolmogorov and Leonid Levin on randomness. In Laurent Mazliak and Glenn Shafer, editors, *The Splendors and Miseries of Martingales: Their History from the Casino to Mathematics*, pages 405–414. Birkhäuser, Cham, 2022.

[4] Gregory J. Chaitin. Information-theoretic limitations of formal systems. *Journal of the Association for Computing Machinery*, 21:403–424, 1974.

[5] Samuel Epstein. All sampling methods produce outliers. *IEEE Transactions on Information Theory*, 67:7568–7578, 2021.

[6] Samuel Epstein and Leonid A. Levin. On sets of high complexity strings. Technical Report arXiv:1107.1458 [cs.CC], arXiv.org e-Print archive, July 2011. Latest version (v7): July 2012; withdrawn in February 2014 and reposted as arXiv:1403.4539 [cs.CC] by Levin.

[7] Péter Gács. On the relation between descriptional complexity and algorithmic probability. *Theoretical Computer Science*, 22:71–93, 1983.

[8] Péter Gács, Yuri Kalnishkan, Alexander Shen, and Vladimir Vovk, editors. *Special Issue of* Information and Computation *in memory of Vladimir V'yugin*, volume 306–312 (depending on the article), 2025–2026.

[9] Péter Gács, John T. Tromp, and Paul Vitányi. Algorithmic statistics. *IEEE Transactions on Information Theory*, 47:2443–2463, 2001.

[10] Leonid A. Levin. Occam bound on lowest complexity of elements. *Annals of Pure and Applied Logic*, 167:897–900, 2016.

[11] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[12] Gleb Novikov. Relations between randomness deficiencies. Technical Report arXiv:1608.08246 [math.LO], arXiv.org e-Print archive, August 2016. Published in *Lecture Notes in Computer Science* 10307:338–350 (2017).

[13] Aaditya Ramdas and Ruodu Wang. Hypothesis testing with e-values. *Foundations and Trends in Statistics*, 1(1–2):1–390, 2025.

[14] Alexey Semenov, Alexander Shen, and Nikolai Vereshchagin. Kolmogorov's last discovery? (Kolmogorov and algorithmic statistics). *Theory of Probability and its Applications*, 68:582–606, 2024. Also published as arXiv:2303.13185 [math.LO], October 2023.

[15] Alexander Shen. The concept of Kolmogorov (α, β)-stochasticity and its properties. *Soviet Mathematics Doklady*, 28:295–299, 1983.

[16] Alexander Shen, Vladimir A. Uspensky, and Nikolai Vereshchagin. *Kolmogorov Complexity and Algorithmic Randomness*. American Mathematical Society, Providence, RI, 2017.

[17] Alexander Shen and Nikolai K. Vereshchagin. *Computable Functions*. American Mathematical Society, Providence, RI, 2003.

[18] Nikolai Vereshchagin and Alexander Shen. Algorithmic statistics: forty years later. In Adam Day, Michael Fellows, Noam Greenberg, Bakhadyr Khoussainov, Alexander Melnikov, and Frances Rosamond, editors, *Computability and Complexity: Essays Dedicated to Rodney G. Downey on the Occasion of His 60th Birthday*, volume 10010 of *Lecture Notes in Computer Science*, pages 669–737. Springer, Cham, 2017.

[19] Vladimir Vovk. Testing exchangeability in the batch mode with e-values and Markov alternatives. *Machine Learning*, 114:article 99, 2025.

[20] Vladimir V. V'yugin. On nonstochastic objects. *Problems of Information Transmission*, 21:77–83, 1985.

[21] Vladimir V. V'yugin. Kolmogorov complexity in the USSR (1975–1982): Isolation and its end. *Information and Computation*, 307:article 105374, 2025. Also published as arXiv:1907.05056 [cs.GL], July 2019.